HEIDELBERG UNIVERSITY
DEPARTMENT OF ECONOMICS

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Forecast revisions in the presence of news: a lab investigation

Joep Lustenhouwer

Isabelle Salle

AWI DISCUSSION PAPER SERIES NO. 714
April 2022

Forecast revisions in the presence of news: a lab investigation

Joep Lustenhouwer* Isabelle Salle†

April 12, 2022

Abstract

We conduct a laboratory experiment in a fully-fledged macroeconomic model where participants receive information about future government spending shocks and are tasked with repeatedly forecasting output over a given horizon. By eliciting several-period-head predictions, we investigate forecast reaction to news and revision. The lab forecasts are consistent with stylized facts on reaction to news established in the survey literature. We find that subjects steadily learn the magnitude of the effect of the shocks on output, albeit not to full extent. We further find little support for fully backward-looking expectations. We rationalize the experimental data in the context of a Bayesian updating model, which provides a better description of the behaviors in longer-horizon environments and among more attentive and experienced subjects.

1 Introduction

We design a laboratory experiment where participants are tasked with repeatedly forecasting output several periods ahead while receiving public information about future government spending shocks. We focus on how forecasts react to such news and we seek to identify the extent and the drivers of forecast revisions.

How agents revise their forecasts and react to news about the economy is certainly relevant, both theoretically and empirically. In standard state-of-the-art models, the expectation channel plays a key role in driving macroeconomic responses to policies and announcements. On the policy front, communication has become a widely popular tool, not only of the central banks when it comes to forward-guidance policies, but also of the fiscal authorities when it comes to counteract negative shocks, such as the initial onset of the COVID-19 pandemic.

While models can provide a great deal of prediction as to how fiscal or monetary policies may influence macroeconomic outcomes, their conclusions hinge to a large

*Heidelberg University, Germany. Financial support for this project by the Deutsche Bundesbank is gratefully acknowledged.

†Bank of Canada, Financial Markets Department, Ottawa, CA & Amsterdam School of Economics, University of Amsterdam, NL & Tinbergen Institute. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Canada or of the Deutsche Bundesbank.

extent on the expectation assumption. Eventually, how expectations are actually formed and successively react to changes in the information set is an empirical question. To investigate it, two routes are available to researchers: the survey route and the lab route. Both approaches have their relative merits.[1] but, while there is a rather large body of evidence on forecast reactions to news and forecast revisions coming from survey studies, the topic has been much less explored using laboratory experiments; see the literature review below.

Using lab evidence is a promising approach to understanding how forecasts react and adjust in the presence of new information. In contrast to surveys, the lab setup offers control over three critical components of decision-making: information, incentives and fundamentals. In a lab experiment, information is controlled via the graphical user interface of the participants and the instructions, which ensures that all have access to the same set of information and this set is public knowledge. In particular, our instructions explain in plain language how government spending and output relate to each other. The interface also explicitly depicts news about future government spending shocks using high-quality graphs.

All participants also share the same objective function, namely the sole minimization of forecast errors, which remove other forms of incentives such as reputational concerns or forecast smoothing. Such control greatly limits the impact of confounding factors that may be brought up to explain deviations from the fully-informed rational expectation (FIRE) benchmark.

Regarding control on fundamentals, the elicited forecasts in the lab are embedded into a fully-fledged macroeconomic model that grossly mimics the dynamics at play in the real world between the main endogenous variables of interest. Despite behind-the-scene tedious algebra that only serves the purpose of micro-founding the underlying model of the experiment, participants' aggregate forecasts end up into a simple, linear, one-dimensional data generating process (DGP). To simplify further, we consider a deterministic model with news about future non-persistent shocks, such that the only noise from the participants' perspective stems from movements in the group's average expectations.[2] This discipline allows us to confront the participant's forecasts and forecast revisions with the perfect-foresight (PF) path and precisely measure deviations from the FIRE model. The controlled environment also ensures that the news are credible and actually materialize.

Within such a controlled environment, we can identify whether deviations from the FIRE model, if any, may be explained by the lack of rational expectations rather than the lack of full information, while such a distinction may at best be

[1]See, e.g., Salle 2022 for a comparative discussion of the two.

[2]On a side note, our model features a weaker-than-usual expectation channel. This means that an increase in the group's expectations tends to increase the actual realization of the predicted variables, but by a lesser extent than usually implemented in the related literature (see Footnote 5 below), where actual realizations are typically mainly expectation-driven. Our claim is that a somewhat attenuated expectation channel is closer to the way naturally-occurring economic systems operate, where strategic uncertainty considerations are not absent but diluted at the aggregate.

suggested using survey data. In other words, the lab experiment isolates behavioral explanations to deviations from the FIRE model.

The design of the four treatments of our experiment is largely exploratory because theory does not indicate which variations of the model may lead to greater or lesser deviations from the FIRE benchmark. The first treatment variable is the length of the horizon, that we denote by T, over which participants have to submit output forecasts. In half of the experimental economies, $T = 2$, such that participants have to submit one-step and two-step-ahead forecasts. In the other half, $T = 6$, such that they have to submit forecasts up to and including six-period ahead. This first treatment dimension varies the cognitive load of the subjects and the extent to which they are forward-looking.[3]

The second treatment variable is the source of funding of the government spending shocks, namely they are either debt- or tax-financed. This second treatment variable results in variations in the time series of the fiscal variables on the interface of the subjects but does not imply any theoretical differences regarding the effect of fiscal shocks on output in the underlying model.

We find three main results. First, in all configurations of the experiment, subjects display learning and the distance between the observed output path and its PF counterpart shrinks over time. While tax or debt-financed government spending does not influence subjects' forecasting behavior, their planning horizon does. In particular, longer-horizon forecasting is associated with smaller deviations from the PF path and better forecasting performances.

Second, we find that the experimental data are qualitatively consistent with the two main stylized facts from the survey literature, namely i) information rigidities and underreaction to news at the aggregate level and ii) overreaction to news in the individual forecasts, which stands in contrast with individual rationality. This result is an addition to the recent efforts to compare survey and experimental forecast data in order to strengthen the external validity of this class of experiments.[4]

Third, the experimental data does not provide strong support for adaptive or extrapolative expectation formation processes, contrary to what most so-called learning-to-forecast experiments (LtFEs) have reported. We conjecture that the provision of information about the future periods and the elicitation of forecasts several period ahead may explain why backward-looking behavior is somewhat dampened in our experiment. We rationalize the experimental data within a Bayesian updating model where agents successively revise their beliefs about the effect of the shocks on output. We find that this model provides a better description of the behaviors in

[3] Evans et al. (2019) find that long-horizon expectations are stabilizing compared to one-step-ahead predictions but their model and the experimental task greatly differ from our setup, which does not allow us to rely on their conclusions to predict treatment effects in the context of the present experiment.

[4] In particular, our results extend the study by Cornand and Hubert (2020) who show that experimental and survey forecast data share several key characteristics and moments, but their experimental dataset does not include data on forecast revisions.

longer-horizon environments and among more attentive and experienced subjects.

The rest of the paper is organized as follows: after a literature review, Section 2 presents the experimental design and implementation, Section 3 gives an overview of the experimental results, Section 4 compares the experimental data with stylized facts on expectation formation in the presence of news established in survey data, Section 5 develops a model of Bayesian updating to rationalize how subjects learn about government spending shocks, and Section 6 concludes.

Literature review There is a relatively vast literature on survey data of macroeconomic variables, whether relying on surveys of professional forecasters, market participants, households or firms. We do not intend to give an exhaustive account of this literature; we rather restrict our attention to the specific study of forecast reactions and revisions in the presence of new information.

The survey literature usually provide evidence that forecasters revise their expectations in an inefficient manner compared to the FIRE benchmark that prescribes the optimal incorporation of news into forecasts. Yet, the results are mixed. For instance, Coibion and Gorodnichenko (2015) and Bordalo et al. (2020) find, respectively, underreaction of aggregate expectations and overreaction of individual expectations to new information. Furthermore, Coibion and Gorodnichenko (2012) find underreaction to different shocks that have been identified in the literature, including the news shock of Barsky and Sims (2011). Using a similar analysis, Angeletos et al. (2021) find that survey expectations initially underreact to a shock innovation, but subsequently overshoot the FIRE path.

Within the survey literature, information-provision experiments in real-world surveys have turned particularly relevant for studying the reaction of expectations to news; see Haaland et al., 2022; Fuster and Zafar, 2022 for an overview. This class of experiments has brought mixed evidence regarding whether the revision of expectations is generally in line with Bayesian updating. For example, Binder and Rodrigue (2018) find that survey participants revise their long-run inflation expectations in the direction of the target after being informed about its value (2%). The authors further find that subjects with higher initial uncertainty about long-run inflation revise their expectations more, which is also consistent with Bayesian updating. Similarly, Armantier et al. (2016) study revisions of inflation expectations after provision of inflation-relevant information and find revisions are proportional to the strength of the information signal, and inversely proportional to the precision of prior inflation expectations. Similar findings for expectations of firms are obtained by Coibion et al. (2018).

However, the extent to which participants display Bayesian updating also appears to be sensitive to the type of information provided. For example, Cavallo et al. (2017) find that participants are more influenced by information that is less costly to understand (such as supermarket prices) relative to information that is more costly to understand (such as inflation statistics). Coibion et al. (2022) find similar differences in expectation revisions across the provision of different types of interest rate information. Moreover, Coibion et al. (2020) find that participants do not revise their expectations when they are provided information about the Fed's new average inflation targeting regime compared to information about regular inflation targeting.

As mentioned in the introduction, the literature on lab experiments investigating forecast revision to news is much more sparse than its survey counterpart. This is probably because lab experiments have been more concerned with one-step-ahead predictions than with the repetitive elicitation of expectations at various, overlapping horizons, as is the case in the present experiment. This is certainly true for the relatively large literature on LtFEs, to which our experiment belongs.[5]

Evans et al. (2019) and Anufriev et al. (2020) constitute two exceptions that are concerned with the elicitation of longer-than-one-horizon predictions. Yet, neither contribution consider the provision of news ahead of the elicitation of forecasts, and neither focus on the process of forecast revision. LtFEs have been designed to be backward-looking, in the sense that subjects are only provided with past time series, possibly coupled with some information about the fundamentals of the economy or the market. We are aware of some recent exceptions where it is investigated how central bank projections of future policy stance or economic developments may influence subjects' inflation and output expectations in a LtFE based on a NK model; see Kryvtsov and Petersen (2015, 2021); Mokhtarzadeh and Petersen (2021); Rholes and Petersen (2021); Ahrens et al. (2022). Yet, subjects in these experiment are tasked with forming one-step-ahead predictions only. The role of the forecast horizon is therefore absent.

In the present paper, we instead build an LtFE in which subjects are given news about the future developments of macroeconomic variables that directly influence the variables that they are tasked with forecasting, and their forecasts span over multiple future periods, so that several vintages co-exist and forecast revisions may be measured.

Finally, parallel to the LtFE literature, individual decision-making experiments have also documented expectation formation, where their self-referential nature is absent; see Afrouzi et al. (2021, Table A.1) for a summary.[6] This literature also

[5]In LtFEs, individual expectations are elicited in a group experiment where the group represents a self-referential system: the subjects' expectations feed back into the realization of the group-level variables that they are tasked to predict in order to reproduce the expectation channel at play in macro-finance models and systems. – see, inter alia, Bao et al. (2021) for a survey of recent developments.

[6]Market experiments, such as Haruvy et al. (2007), also elicit future price expectations but forecasts are not the central focus of these studies; they are collected to enhance the understanding of the trading behaviors of the participants. We will also refrain from discussing the large literature in behavioral finance on the role of information and news on the occurrence of bubbles; for instance, Marquardt et al. (2019) show how news are rapidly integrated into stock prices in an experimental asset market where participants are traders.

supports deviations from the FIRE benchmark, in the form of various adaptive or extrapolative models. In this class of experiments, the focus has equally been on forecast reaction to contemporaneous shocks or adaptation to past trends. Subjects are typically not projected with future values of variables that are relevant to form their expectations.

As regards to the fiscal policy set-up of our experimental environment, our paper is related to Meissner and Rostam-Afschar (2017) who test whether subjects learn to comply with Ricardian Equivalence in a laboratory experiment. However, their experiment features a life-cycle consumption environment where subjects need to make consumption decisions and are not explicitly asked to make forecasts. Our focus on forecasting and revisions of expectations in light of shocks thus provides a contribution that is quite different from theirs and from other macro-finance experiments on consumption allocation, price setting, investing or other economic decisions.

2 The experiment

We first introduce the underlying model that provides the DGP of the experimental economies before explaining our design and lab implementation.

2.1 The underlying model

We use a standard dynamic stochastic general equilibrium (DSGE) model with sticky prices, an inflation-targeting interest rate rule and a government sector. The details of the underlying model are deferred to Appendix A. Even through the model architecture may look complicated at a first glance, its only purpose is to serve as a structured framework in which to elicit expectations at various horizons in an economically meaningful context in the laboratory. We wish to reassure the reader that the instructions and the experimental task of the participants are simple enough to avoid confusion.

Before diving into the experimental design, some comments are in order regarding the implementation of finite planning horizons in the model. The microfoundations of the DSGE models require agents to solve an infinite-horizon planning optimization problem and, hence, form infinite-horizon expectations of the relevant economic variables. To use the reduced-form model that features only one-step-ahead expectations under learning or in the LtFE literature, one must assume so-called Euler-equation learning for the recursive formulation of the decision problem to be valid under non-RE; see, e.g., Evans and Honkapohja (2012). However, since we need a model with longer-than-one-horizon expectations, we choose an alternative form of bounded rationality, under which agents have finite planning horizons along the lines of Woodford (2019); Lustenhouwer (2020); Lustenhouwer and Mavromatis (2021).[7] Our implementation enables us to derive a data-generating process (DGP) for any arbitrary finite planning horizon T, which makes straight-forward the design of treatments with distinct horizon lengths.

To be precise, the DGP underlying the experimental economies for any finite planning horizon T reads as follows:

$$(\tilde{\rho} - \nu_y)\hat{Y}_t = \tag{1}$$
$$\tilde{\rho}\bar{g}\hat{G}_t + \tilde{\nu}_g \sum_{s=0}^{T} \beta^s (\bar{E}_t \hat{G}_{t+s}) + \nu_y \sum_{s=1}^{T} \beta^s (\bar{E}_t \hat{Y}_{t+s}) - \mu \sum_{s=1}^{T} \beta^s \sum_{j=0}^{s-1} (\bar{E}_t i_{t+j} - \bar{E}_t \pi_{t+j+1}),$$

which corresponds to Equation (72) in Appendix A, where $\bar{g}$ denotes the steady state value of government spending as a fraction of output, and the composite parameters $\mu, \nu_y, \tilde{\rho}$ and $\tilde{\nu}_g$ are given by, respectively, Equations (61),(66),(73) and (74) in that same appendix.

In the experiment, we set the T-period-ahead path of government spending $\{\hat{G}_{t+s}\}$ and we only elicit T output expectations $\{\bar{E}_t \hat{Y}_{t+s}\}_{s=0,...,T}$, as explained in Section 2.2 hereafter. Hence, to solve for contemporaneous output $\hat{Y}_t$ on the LHS of (1), we need to pin down the remaining ingredient on the RHS, namely the T-period-ahead expected real rate path $(\bar{E}_t i_{t+j} - \bar{E}_t \pi_{t+j+1})$, which encompasses expected inflation rates $\bar{E}_t \pi_{t+j+1}$ and expected nominal interest rates $\bar{E}_t i_{t+j}$.

To keep the cognitive load and the complexity of the forecasting tasks manageable for the subjects, we assume that the aggregate expectations of inflation and the nominal interest rate are consistent with the subjects' elicited output expectations, as described in Appendix A. Specifically, inflation expectations are computed per Equation (79), while nominal rate expectations are computed per Equation (80), where both formulas only depend on the expected output values elicited from the subjects and the planned government spending set by the experimenter.

Automatizing real rate expectations further allows us to simplify the presentation of the relatively complex machinery behind the macroeconomic model in the experimental economy. Subjects could then fully focus on a relatively simple setup where output and government spending are the variables of interest and inflation and rates are kept implicit. Furthermore, choosing real rate expectations to be consistent with output expectations ensures that deviations of the human-generated output expectations are the only potential source of deviations from the PF benchmark.[8]

This model provides the necessary framework in which multiple-step-ahead fore-

[7]Our approach is similar to Evans et al. (2019) insofar as bounded rationality is integrated in the data-generating process (DGP) for the purpose of designing an underlying model to an LtFE.

[8]Under our calibration, see Table 7 in Appendix A, our specification of the fiscal shocks and our choice of the values of the treatment variable T, see Section 2.2, the PF equilibrium output path of the experimental economies is given by $Y_t = f(G_t, G_{t+1}, ..., G_{t+6}) = 0.243G_t - 0.009G_{t+1} - 0.010G_{t+2} - 0.011G_{t+3} - 0.012G_{t+4} - 0.012G_{t+5} - 0.013G_{t+6}$ in the case where $T = 6$ and $Y_t = f(G_t, G_{t+1}, G_{t+2}) = 0.243G_t - 0.009G_{t+1} - 0.010G_{t+2}$ in the case where $T = 2$.

casts can be elicited and confronted with a FIRE benchmark. We now detail how the experiment works.

2.2 Experimental design

Subjects are assigned to groups of 10. The composition of the group does not change throughout the experimental session.

Summary of the treatments We consider a 2-by-2 design where we vary T, the horizon of the forecasts, and the way public spending shocks are financed.

Regarding the forecast horizon, we set to $T = 2$ or $T = 6$, so that participants in groups with $T = 2$ have to forecast output for $t + 1$ and $t + 2$ (i.e. $T = 2$ in Eq. (1)), while those in groups with $T = 6$ have to submit six predictions per period, i.e. one prediction for each future period between $t + 1$ and $t + 6$ (i.e. $T = 6$ in Eq. (1)). As for the second treatment variable, the public spending shocks are either tax-financed or debt-financed.

We explain below what the implied differences in terms of the model are as well as the experimental task in each of the four resulting treatments: $T = 2$ and tax-financed shocks, $T = 6$ and tax-financed shocks, $T = 2$ and debt-financed shocks and $T = 6$ and debt-financed shocks.

Experimental task The duration of the experiment is 80 periods. In each period, subjects' experimental task is to forecast the level of output for all future periods within their planning horizon T. The level of output, as well as all other macroeconomic variables in the experiment are expressed as percentage points (p.p.) deviation from their steady-state values. The planning horizon remains fixed for the full length of the experiment and is the same for every subject in a given group. Subjects were aware of these elements.

Appendix C provides the exhaustive lab material, including the instructions, the graphical user interface (GUI) for each horizon treatment, the quiz and the post-experiment questionnaire.

Payoff Subjects are paid according to their forecast accuracy, as usual in LtFEs. Not every single submitted forecast counts towards their earnings. One vintage is randomly drawn with equal probability in each period to be rewarded. More precisely, each vintage is rewarded with probability 1/6 in the $T = 6$ treatment and 1/2 in the $T = 2$ treatments. The same vintage series is used for all the subjects within a given group to ensure fairness.

The payoff function is decreasing with the absolute forecast error of the randomly drawn vintage. Subjects earn points in each period according to 100/(1 + absolute forecast error). At the end of the experiment, subjects have earned a cumulative score across the 80 periods[9] that is exchanged against euros at the rate of 300 points for 1 Euro. The instructions include a complete description of the payoff scheme.

Timing of events with an experimental period Within each period $t = 1, \ldots, 80$, events unfold as follows. Once every subject in the group has submitted their forecasts for the next T periods, the aggregate output forecasts are computed as the arithmetic means of the ten individual forecasts of the group. This is done separately for each of the future T periods. One then obtains a time series of length T of aggregate output forecasts, from the horizon $t+1$ up until horizon $t+T$, that is inserted into the DGP (1) in lieu of $\{Y_{t+\tau}\}$, $\tau = 1, \ldots, T$, where T is either 2 or 6 depending on the treatment.

To solve for contemporaneous output in (1), the remaining required ingredient is the path of the future public spending shocks G for the next T periods. This path is *ex ante* and exogenously set and is the same for every group in the experiment. On their GUI, subjects observe the upcoming government expenses for their forecasting horizon, namely for the next T periods; see Figure 1 and details in the next paragraph below. This time series $\{G_{t+\tau}\}$, $\tau = 1, \ldots, T$ is plugged into the DGP (1) which then, together with the aggregate output forecast paths, delivers the realization of output in period t. The experimental game then moves to the next period and unfolds this way for 80 periods.

Government expenditure shocks We consider positive and negative shocks and various magnitude to assess whether participants may learn the fundamental relationship between government spending and output. Specifically, we use the following seven shocks: government expenditures equal 0 in every period except in period 17, where $G(17) = 15$, in period 27, where $G(27) = -10$, in period 36, $G(36) = 20$, in period 44, $G(44) = 10$, in period 54, $G(54) = 17.5$, in period 63, $G(63) = -15$ and in period 73, $G(73) = -20$.

In the model, each such shock is a news arising within the planning horizon of the agents; see below the description of the information set of the subjects in each period. As usual in experimental macroeconomics, the magnitude of the shocks is relatively larger than their empirical counterparts to make them salient enough to the subjects.

We use one-time, i.e. non-persistent shocks, for several reasons. A simple shock time series simplifies the economic environment in which the subjects have to submit their forecasts. It may be necessary given the quantity of information already present on the GUI; see Figure 1 below and Appendix C. With such a stylized shock structure, we can straight-forwardly assess whether subjects learn how to optimally

[9] Notice that forecasts submitted for periods $80+1$ to $80+T$ are not incentivized and, hence, not included in our dataset, because the realizations corresponding to the submitted forecasts never materialize (since the experiment ends in period 80 and subjects are aware of this).

respond to announcements that have an easy relation to the variable they forecast (namely output).

Moreover, we design the shock series such that non-zero spending shocks are at least one planning horizon apart from each other. One-time shocks, together with enough distance between these, make sure that subjects only see one upcoming shock at a time, which allows for a clean identification of which news subjects react to. By contrast, if shocks were to be auto-correlated, subjects would incorporate the effects of multiple shocks simultaneously into their forecasts because the realizations of multiple shocks would unfold within the same T-period horizon.[10] We would then no longer be capable of identifying the forecast responses and revisions to a particular piece of information. This is because behavioral expectations are sluggish: while under the FIRE model each news is immediately incorporated into the relevant forecasts, this cannot be postulated *ex ante* in the lab. It may take several periods for subjects to adjust their forecasts even if no new information is revealed across time.

Finally, non-persistent shocks also allows us to attribute any persistence in output following the shocks as stemming from the expectations of the subjects only. We may then conveniently isolate the contribution of possible deviations from the FIRE benchmark on output deviation from the PF equilibrium path.

Information and graphical user interface Participants do not know the exact form and the coefficient values of the DGP. It should be argued that the choice of not spelling out Equation (1) in the instructions finds at least two rationales. We believe that communicating mathematical expressions to non-experts such as students from various fields is likely to increase rather than mitigate confusion. We rather devote considerable efforts in the instructions to explain in plain language the relationships between the main variables of the model, namely government spending and output; see, again, Appendix C.[11]

In the context of our experiment, grasping Eq. (1) would arguably involve a large cognitive load. Yet, this level of complexity is not an unrealistic feature of our setting because in the real world, the true DGP of the economic system is certainly complicated and equally unknown. Therefore, a second rationale for keeping the DGP away from the subjects is to mimic the level of knowledge that actual agents are endowed with in naturally-occurring environments.

Figure 1 reproduces the GUI of the subjects in the treatments where $T = 6$.[12] As

[10]For example, consider the case of $T = 6$, where the news in period 11 is that $G(17) \neq 0$. Then, in period 12, in the case of an auto-correlated shock, subjects would not only see that $G(17) \neq 0$, but also that $G(18) \neq 0$. What is more, when period 21 comes, subjects receive the news about the second shock, namely $G(27) \neq 0$, while still observing that $G(22), ..., G(26) \neq 0$ due to the persistence of the initial shock in period 17.

[11]In the instructions, we also explain the intuitive role of private consumption in mitigating the effect of government spending on output.

[12]We refer to Appendix C for the complete set of lab materials for each treatment.

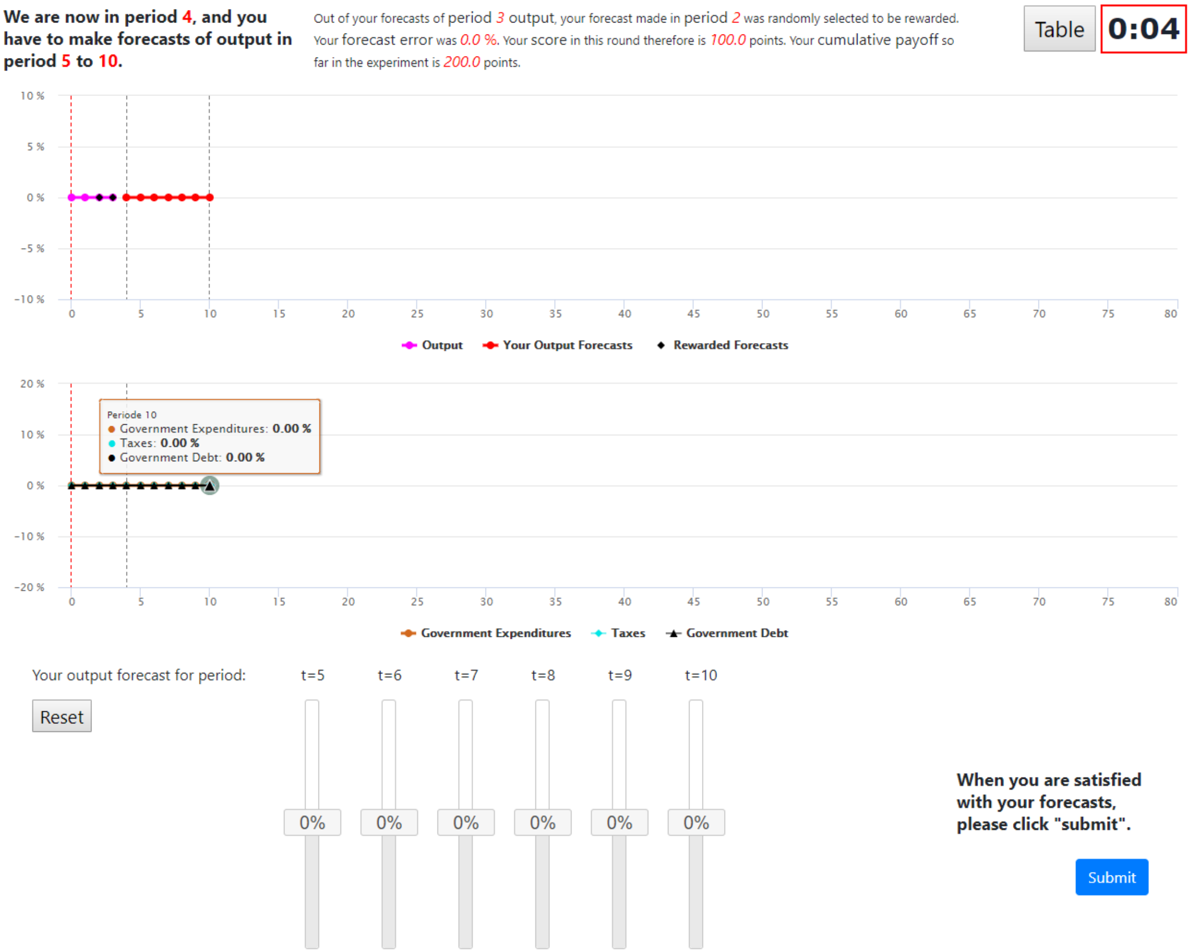
one can see, to submit their forecasts, subjects are presented with sliders that they could drag – in steps of 0.1 unit – to the value they wish to submit. There is one slider for each forecast, so six sliders when $T = 6$ and two when $T = 2$. The range of the sliders is chosen to allow for a relatively wide range of values but to rule out wild values, so that it is not possible for subjects to try to create extreme fluctuations in the artificial economy by submitting forecasts that are orders of magnitudes larger than the realized values under the PF benchmark.

It is important to realize that expectation revision is a large part of the experimental task because subjects submit multiple vintages. For instance, take the case where $T = 6$. Upon entering the experimental game, subjects submit output expectations for periods 2, 3, 4, 5, 6 and 7. In the next period – period 2, subjects have to forecast again output for periods 3 to 7 and add a new forecast, for period 8, etc. Hence, they revise multiple times their forecasts.

To facilitate the revision process, at the beginning of each period, we place each slider at the value corresponding to their latest forecast for that given period. This way, subjects do not have to recall their previous forecasts when revising their forecasts in subsequent periods. As for the slider corresponding to the new period's forecast, namely period $t + T$, for which no previous forecast has been submitted, it is initialized at a neutral value of 0, which corresponds to the steady state level of output.

As can be seen in Fig. 1, the GUI provides each subject with information about their own forecasts. For each past period, the forecasts that were chosen to count towards earnings were displayed as a dot on the output graph. For each upcoming period within the T-period-ahead horizon, their most recent forecasts about the respective periods are projected in the graph, and these are simultaneously adjusted as subjects move their sliders.

The GUI also displays the evolution of each macroeconomic variable in graphs and in a table. Output is displayed up until the previous period. The spending shocks, that constitute the news, along with the corresponding levels of taxes and government debt are announced T periods in advance; see the next paragraph for more detail. Subjects are told in the instructions that the government always sticks to the announced path and, indeed, they are never revised. Therefore, subjects become aware of the fiscal shocks once their period of realization is included into their horizon. For instance, in treatments with $T = 6$, subjects see the first shock (occurring in period 17) for the first time in period $t = 11$, when they have to forecast output over periods 12 to 17 and they then see government spending, taxes and debt up until period 17. In the table, subjects also explicitly see the absolute forecast errors that count towards their earnings and their resulting points in each period.



Notes: For completeness, the GUI in experiments with $H = 2$ is reported in Appendix C.

Figure 1: Example of the graphical user interface (GUI) with $H = 6$

Tax- versus debt-financed government spending The GUI is where the second treatment dimension, namely whether public spending shocks are tax- or debt-financed, becomes salient. Taxes are taken to be lump sum. In the tax-financed treatment, taxes always move one-to-one with government spending, whereas debt remains at zero in all periods. In these treatments, the graphs of government spending and taxes hence perfectly overlap and information about taxes is redundant.

In the debt-financed treatments, taxes remain at 0 throughout the experiment, whereas debt evolves according to the government budget constraint as dictated by (69) in Appendix A:

$$\tilde{b}_{t+1} = \frac{\bar{G}}{\bar{Y}\beta}\hat{G}_t + \frac{1}{\beta}\tilde{b}_t. \tag{2}$$

Note that, since the experiment features both positive and negative spending shocks, government spending can be seen as the instrument that stabilizes debt in the long run in the debt-financed treatments.

Importantly, in the model, output does not depend on how government spending

Financing of the fiscal shocks	tax-financed	debt-financed
Horizon T of the forecasts		
$T = 2$	4	4
$T = 6$	4	4

Table 1: Number of independent observations per treatment

is financed. In other words, Ricardian equivalence holds, as expectations beyond the horizon T are assumed to be formed accordingly (see Appendix A). In the instructions, subjects are explicitly told about this piece of information in all treatments, and subjects do not know prior to starting the game whether their experimental economy features only debt- or only tax-financed spending. However, subjects are also told that the way government spending is financed may affect the dynamics of aggregate output in so far as it may affect the expectations of the participants in their group.

2.3 Lab implementation

The experiment is programmed within the `oTree` experimental software of Chen et al. (2016). Sessions took place between March and July 2019. In total, we conducted four groups with ten participants for each treatment so that a total of 160 subjects participate in our experiment. Table 1 summarizes the number of independent observations per treatment. Half of the sessions/groups within each treatment were conducted in the CREED laboratory of the University of Amsterdam (Netherlands) and the other half were conducted at the University of Bamberg (Germany).[13] The sessions at the University of Amsterdam use instructions and a GUI in English, as usual at the CREED lab. As for the sessions in Bamberg, we translated the instructions and the GUI to German; see Appendix C.

Subjects had to correctly answer a quiz before starting the experiment, possibly with the guidance of the experimenter. Only once every subject in a group had done so could the group proceed to the experimental game. This ensures that the information contained in the instructions are common knowledge. At the end of the experiment, subjects were asked to answer a questionnaire, containing in particular demographics.

Since our experiment is a group experiment subjects can only move to the next period once all 10 subjects have submitted their forecasts. In order to motivate subjects to submit their forecasts within a reasonable time-frame not to delay the experiment, a count-down timer is displayed at the top of their screens. Once this timer reaches zero, a pop-up message appears with the text“Time is up!”. This timer is not binding, and subjects can still proceed to submit their forecasts after the message appeared.

[13]To be precise, within each treatment, two groups were run in Amsterdam and two in Bamberg.

Including the reading of instructions, the quiz and the end questionnaire, an experimental session lasts for less than two hours and the average earnings are 19.0 euros (with a standard deviation of 2.2 Euros). All subjects are students (partly in economics but also in other fields) and no subject participates more than once in the experiment. Some subjects have prior experience with LtFEs, but more than half do not.

We now turn to the experimental results.

3 Main insights on expectation formation and revision

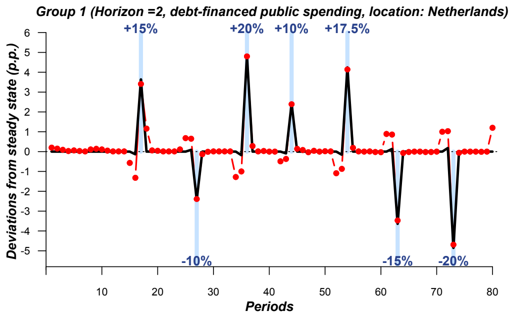
In this section, we highlight the main determinants of group dynamics and forecast revisions observed in the experiment.

3.1 Aggregate cross-treatment dynamics

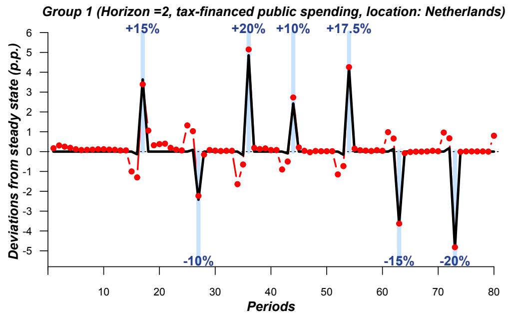
Figure 2 exemplifies one experimental economy in each of the four treatments by plotting the aggregate realized output next to its PF path and the fiscal shocks implemented in the experiment (we recall that all values are expressed in deviation from their steady state values in p.p.). Figures 6-9 in Appendix B report all 16 groups.

A first glance at the data shows that output deviates from its PF path. Because the environment is otherwise deterministic, these deviations are entirely due to deviations of the output forecasts from PF. Admittedly, this is not unexpected: only under the FIRE assumption, in particular under the assumption of full information regarding the DGP, along with common knowledge, may we observe no deviation from the PF path, and zero forecast errors, right from the start of the experiment. More interesting are the following two sets of questions: i) are these deviations from the PF path treatment-dependent?; ii) do these deviations fade away over time; in other words do participants *learn* to form forecasts that are closer to PF?

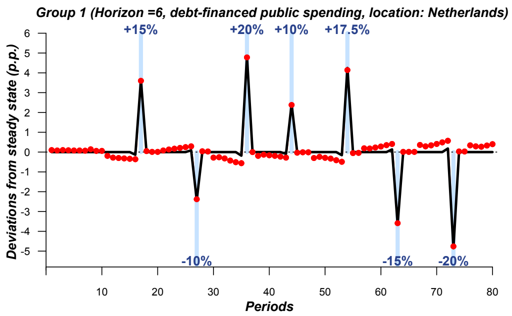
Regarding the first question, eye-balling Figure 2 reveals that deviations from PF may be partly treatment-dependent: deviations from PF before the shocks appear greater in the treatments where participants form up to two-period-ahead expectations ($T = 2$) than up to six ($T = 6$) – look at the distance between the red dots and the solid black lines before each shaded area. Output deviations from the PF equilibrium also appear larger in the aftermath of the shocks with $T = 2$ than with $T = 6$. Whether the shocks are tax- or debt-financed does not seem to influence the output paths – compare the left-hand-side and the right-hand-side panels. In what follows, we formally address these two sets of questions by using aggregate data and individual forecast times series.



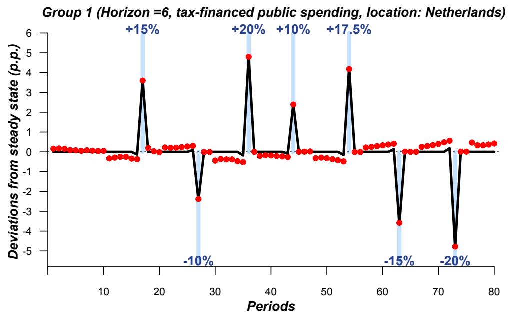
(a) Group 1 (NL), Tr. `H2-debt`



(b) Group 1 (NL), Tr. `H2-tax`



(c) Group 1 (NL), Tr. `H6-debt`



(d) Group 1 (NL), Tr. `H6-tax`

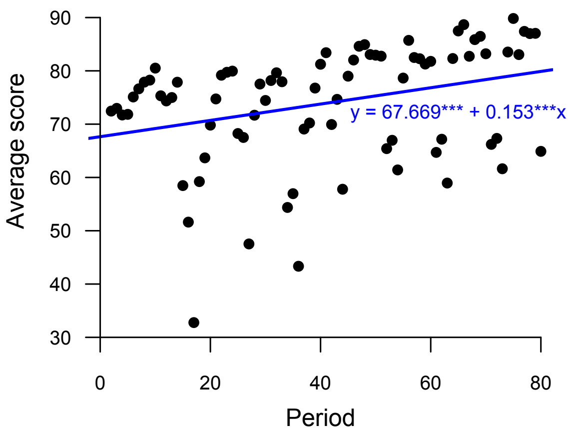
Notes: Each graphic represents the output in the experimental economy (red line with dots), the output under perfect foresight (solid black line), the frequency of each fiscal shock (gray-shaded area) and their size. (File aggData.R up to line 290).

Figure 2: Examples of output dynamics in each treatment

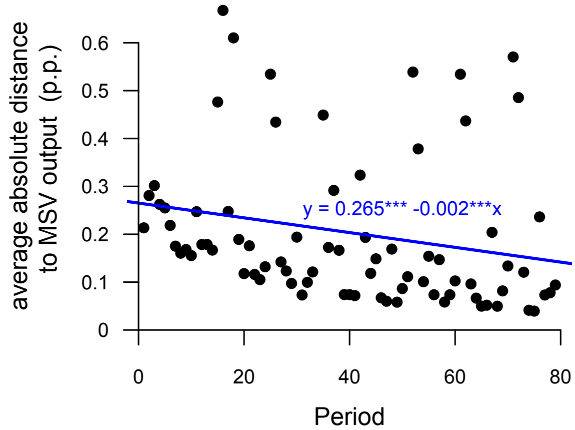
We use two indicators to assess the cross-treatment differences in learning and convergence in the experiment: i) the average earnings efficiency ratios (EERs), that is, the ratio between the average amount of points earned by all the subjects and the total amount of points available in case of zero forecast errors, and ii) the cumulative squared distance of realized output to its PF counterpart, averaged across all groups within a given treatment.

Figure 3 reports the evolution of these two indicators over time across the 80 periods of the experiment, averaged per period over all treatments. It is striking that EERs significantly improve over time, which shows that participants tend to improve their forecasting performances and, hence, their earnings significantly increase throughout the experiment. This is a clear sign that the repeated nature of the forecasting task in the experiment allows learning to take place. As for the deviations from PF output, in line with the learning dynamics just discussed, they significantly decrease over time, which indicates convergence towards the PF path; see Figure 3b.

We now uncover cross-treatment differences in these dynamics with rank-sum tests at the matching-group level; see Table 2. The horizon does not influence the EER (p-value = 0.573) but matters for the distance to PF, where treatments with $T = 6$ result in closer output to its PF path than treatments with $T = 2$ (p-value $< 0.01^{***}$). Neither the location of the experiment, which is reassuring, nor the



(a) Earning efficiency ratios (EERs) per period



(b) Absolute distance to the PF path

Notes: Data are averaged per period over all treatments (i.e. the 16 experimental economies)

Figure 3: Convergence and learning in the experiment

financing of the fiscal shocks (tax or debt-financed) matters for any of these two indicators (the minimum p-value is 0.161).

The lack of cross-horizon differences in the EERs may be explained by two opposing forces.[14] On the one hand, longer-horizon environments are stabilizing, as testified by the lower distances to the PF path than under shorter horizons, which makes forecasting easier and EERs larger in treatments with $H = 6$ than with $H = 2$. This is partly due to the weaker expectation feedback in the treatment with $T = 6$ than $T = 2$.[15] On the other hand, forecasting several-period-ahead is more challenging than forecasting over a shorter horizon because the longer the horizon, the greater the uncertainty. Hence, this second mechanism tends to make EERs smaller with $H = 6$ than with $H = 2$.

We now analyze individual forecast errors.

3.2 Determinants of forecast errors

Our experiment allows us to collect an impressive amount of 48,000 elicited forecast points. Using this data, Table 3 reports the results of pooled OLS regression models where the dependent variable is the squared forecast errors of the participants.[16] The explanatory variables include the available pieces of information at the time of forecasting, the treatment variables and the individual characteristics of each

[14] Evans et al. (2019) discuss a similar mechanism.

[15] Precisely, the expectation feedback is equal to 0.44 when $T = 6$ versus 0.84 when $T = 2$. To obtain these numbers, we consider by how much output in t would react to a uniform increase in expectations for all horizons $t + 1$ to $T + H$. If only one-period-ahead expectations rise by one unit, these values become 0.42 for $T = 2$ versus only 0.08 for $T = 6$.

[16] To be precise, we denote as $\left(e_{t,\tau}^{i,g}\right) = \left(\hat{E}_t^{i,g}(Y_{t+\tau}) - Y_{t+\tau}\right)^2$ the squared forecast error that participant i in group g made in period t when forecasting output in period $t + \tau$, with $\tau = 1, \ldots, T$. The lower the forecast errors, the higher the participant's payoff.

Panel A. Earnings Efficiency Ratios (EER)

	$T=2$	$T=6$	*debt*	*tax*	Amsterdam	Bamberg
Mean	0.724	0.736	0.717	0.743	0.741	0.719
(*sd*)	(0.074)	(0.091)	(0.088)	(0.076)	(0.076)	(0.088)
p-value	0.573		0.161		0.235	

Panel B. Total squared distance to MSV output

	$T=2$	$T=6$	*debt*	*tax*	Amsterdam	Bamberg
Mean	16.664	2.219	9.232	9.651	9.404	9.479
(*sd*)	(5.254)	(0.468)	(7.924)	(9.174)	(8.660)	(8.489)
p-value	$< 0.01^{***}$		0.959		0.645	

Notes: $***$: significant at the 1% level, **: significant at the 5% level, and *: significant at the 10% level. Standard deviations are reported in parentheses below the means. All tests are exact two-sided Mann-Whitney rank sum tests. The tests are performed on matching group level with 8 observations in each group.

Table 2: Output distance to MSV solution and forecasting performances

participant as collected in the end questionnaire.

Col (I) brings three main insights. First, there are clear cross-treatment differences in the forecasting abilities of subjects. Their forecast errors are lower in treatments where they have to forecast up to six-period ahead than two-period ahead. By contrast, the effect of the shocks being tax rather than debt-financed is not robust to the introduction of individual characteristics (Col IV).

The expectation channel being milder with $T = 6$ than with $T = 2$, fluctuations in output are milder and forecasting becomes easier, which leads to lower forecast errors. Another explanation for the cross-horizon differences in forecasting performances could be that participants submit more forecasts (in fact three times as many) under $T = 6$ than under $T = 2$. Hence, experience and training could explain why they become better forecasters than the short-horizon participants. These explanations need not, of course, be mutually exclusive.

Overall, as stressed in Section 3.1, learning occurs in the experiment: forecast errors tend to decrease over time. Finally, in line with intuition and survey evidence, forecasting at a longer horizon entails larger forecast errors. It is also reassuring to see that the location of the experiment does not influence the performances of the participants. These effects are strongly robust across all specifications.

The evidence collected so far leads us to the first insight:

Finding 1 (Learning and convergence to the PF path) *In all treatments, subjects display learning. Longer-horizon forecasts result in more stable output dynamics that are closer to the PF path than shorter-horizon forecasts. Tax- or debt-financing does not influence learning and convergence.*

Col (II) of Table 3 reports on the influence of the spending shocks on forecast errors. Overall, errors are not independent from the shocks, as would be the case under FIRE: the larger the shocks (in absolute value), the larger the forecasting

Dependent variable: Squared forecast errors $\left(e_{t,\tau}^{i,g}\right)^2 = \left(\hat{E}_t^{i,g}(Y_{t+\tau}) - Y_{t+\tau}\right)^2$				
	(I)	(II)	(III)	(IV)
Dummy = 1 for $T = 6$	-0.090^{***}	-0.090^{***}	-0.084^{***}	-0.088^{***}
	(0.017)	(0.019)	(0.019)	(0.017)
Dummy = 1 for tax-financed	-0.040^{**}	-0.042^{**}	-0.039^{**}	-0.032
	(0.017)	(0.018)	(0.019)	(0.024)
Dummy = 1 for location Bamberg	0.014	0.011	0.004	0.014
	(0.014)	(0.015)	(0.017)	(0.017)
Horizon τ	0.010^{***}	0.009^{***}	0.012^{***}	0.012^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Period t	-0.003^{***}	-0.003^{***}	-0.003^{***}	-0.003^{***}
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Announced (squared) shock $G_{t+\tau}^2$		0.0004^{***}	0.0003^{***}	0.0004^{***}
		(0.0001)	(0.0001)	(0.0001)
Past (squared) shock G_{t-1}^2		0.0001^{*}	0.0001^{**}	0.0001^{**}
		(0.00005)	(0.00005)	(0.0001)
Dummy = 1 if negative shock $G_{t+\tau} < 0$		0.003	0.003	0.004
		(0.015)	(0.014)	(0.014)
Dummy = 1 if $\exists \tau \in \{1, \ldots, T\}$ such that $G_{t+\tau} \neq 0$		-0.011	-0.002	-0.002
		(0.016)	(0.015)	(0.015)
Last squared error $(e_{t-1,\tau+1}^{i,g})^2$			0.031^{***}	0.029^{***}
			(0.007)	(0.007)
Time to complete the quiz (cognitive ability)				0.003^{**}
				(0.001)
Dummy = 1 for experience				-0.046^{**}
				(0.020)
Q1 (effort level)				-0.010
				(0.013)
Q2 (immediate reaction to spending shocks)				0.005
				(0.014)
Q3 (delayed reaction to spending shocks)				0.021^{**}
				(0.010)
Q4 (ignoring spending shocks)				0.020^{**}
				(0.009)
Q5 (ignoring government debt)				-0.007
				(0.005)
Q6 (ignoring taxes)				0.002
				(0.005)
Q7 (high cognitive load Q7 (high cognitive load				0.028^{***}
				(0.010)
Constant	0.328^{***}	0.327^{***}	0.300^{***}	0.193^{***}
	(0.020)	(0.020)	(0.017)	(0.046)
Demographics	Yes	Yes	Yes	Yes
# of observations	41,563	41,020	30,967	30,967
R^2	0.041	0.047	0.085	0.110

Notes: *p<0.1; **p<0.05; ***p<0.01. Standard errors are reported below the coefficients in parentheses. Errors are clustered at the group level. Forecasts submitted during the last T periods are non-incentivized, hence discarded. Demographics corresponds to the age, the field of study, the geographical origin and the gender. None of these variables are significant (the lowest p-value is 0.115). The exact questions and their interpretation are given in Appendix C.1. Self-reported answers range from 1 (do not agree at all) to 5 (fully agree).

Table 3: OLS pooled regressions on individual squared forecast errors

errors. Moreover, forecasting errors are larger in the aftermath of a shock (see the coefficient associated to G_{t-1}), but whether the information set of the subjects includes a non-zero spending value does not influence their forecast errors (see the coefficient associated to the dummy variable $G_{t+\tau} \neq 0$). Nor does the sign of the shocks, which speaks against asymmetry responses to shocks (see the coefficient associated to the dummy variable $G_{t+\tau} < 0$). These effects are robust to the introduction of the lagged forecast errors (corresponding to the same prediction period) in Col (III). The coefficient on these lagged forecast errors is positive and significant, implying that forecasts are not optimally revised to correct past mistakes. In the sequel, we further discuss this point.

Finally, Col (IV) shows which individual characteristics of the participants significantly correlate their performances in the experiment. In particular, having previous experiment as a subject in a LtFE improves forecast accuracy. Furthermore, we use the time subjects needed to complete the pre-experiment quiz as a proxy for cognitive ability – the less time they need, the higher their level of reflection. In line with intuition, more skilled subjects forecast more accurately. The answer to Q7 also correlates with performances: the more confused subjects report to be with respect to their experimental tasks, the larger their forecast errors.[17] Answers to Q3 and Q4 also show that subjects who fail to immediately incorporate the information about future spending shocks suffered larger errors.

Before modeling the individual forecasting behaviors, we compare the lab data with the evidence from the survey literature.

4 Experimental versus survey data on forecast revisions

We now confront our experimental data with two well-known stylized facts from the survey literature, namely underreaction to news at the aggregate level and overreaction at the individual level. We show that our data are consistent these two facts.

4.1 Under-reaction to news at the aggregate level

Popular models that incorporate deviations from FIRE include sticky information and imperfect information models. These models imply a positive relationship between the *ex post* mean forecast errors and the *ex ante* mean forecast revisions, where the coefficient on forecast revisions maps one-to-one into the underlying degree of informational rigidities. Such a positive relationship has also been documented in survey data; see, *inter alia*, Coibion and Gorodnichenko (2015).

[17]The list of questions is reported in Appendix C.1.

Dependent variable: aggregate *ex post* forecast errors (Eq. 3)

	All economies ($T=2$ and $T=6$)			Economies with $T=6$ only		
	$\tau=1$			$\tau=3$		
	OLS	FE	FE	OLS	FE	FE
$\hat{b}$	0.134***	0.134***	0.134***	0.058***	0.059***	0.058***
(sd)	(0.028)	(0.017)	(0.017)	(0.012)	(0.014)	(0.014)
y_{t-2}			-0.001			-0.014
(sd)			(0.017)			(0.019)
# economies with sign. $b>0$	9	9	9	5	5	5
# economies	16	16	16	8	8	8

Notes: ***: significant at the 1% level, **: significant at the 5% level, and *: significant at the 10% level. Standard deviations are reported below in parentheses. For OLS regressions in the pooled panel, errors are clustered at the group level. FE refers to panel regressions with fixed effects. For the OLS regressions for each independent group, we use the heteroskedasticity and autocorrelation consistent estimate of the R `sandwich` package.

Table 4: Evidence of under-reaction to news in the experimental economies

To test it in our dataset, we estimate the following relationship for each horizon τ separately:

$$Y_{g,t+\tau} - \bar{E}_{g,t}(Y_{g,t+\tau}) = c + b\left[\bar{E}_{g,t}(Y_{g,t+\tau}) - \bar{E}_{g,t-1}(Y_{g,t+\tau})\right] + \text{error}_{g,t} \qquad (3)$$

where the dependent variable is the *average* forecast error across all 10 participants of a given group g related to the output forecast made in period t for output τ-period ahead in that group – $\tau = 1, \ldots, 5$ with $T = 6$ and $\tau = 1$ with $T = 2$; and the RHS variable is the *average* forecast revision between period $t-1$ and t of the τ-period-ahead forecast in group g.

The estimate of b should be significantly positive for our data to imply information rigidities. Eq. (3) has been developed for single, country-specific survey datasets, while our data encompasses 16 independent experimental economies. Hence, we estimate Eq. (3) first by pooling all the economies together and report the estimate of b for each horizon. We then report the number of economies for which the coefficient b is significantly positive when Eq. (3) is estimated for each of the 16 economies separately. Results are insensitive to either alternative.

Table 4 reports the results for one-step-ahead forecasts (first three columns) where the two treatments with $T = 2$ and $T = 6$ can be included, and the three-step-ahead forecasts in the eight experimental economies with $T = 6$ (last three columns).[18] The results for the other horizons are virtually the same and are deferred to Panel A of Table 8 in Appendix B.

In a nutshell, our experimental data features under-reaction at the aggregate

[18]Because the RHS of Regression (3) requires the lagged $\tau + 1$-ahead-forecast, we may only include one-step-ahead forecasts in the LHS when $T = 2$. Hence, for $\tau = 1$, we have a panel with 16 groups and for $\tau = 2, \ldots, 5$, 8 groups, corresponding to the $T = 6$-treatments.

level and are therefore consistent with information rigidities. This result is robust to considering OLS regressions with clustered errors at the group level or panel models with fixed effects for each group (see the first two columns of each horizon). The last two rows of Table 4 also show that most experimental economies display significant information ridigities when Model (3) is estimated at the group level.

We may then use the results from the one-step-ahead errors ($\tau = 1$) to look into cross-treatment differences along the horizon dimension. We do not find any difference in the estimated group-specific coefficients b: the p-value of the Wilcoxon rank-sum test at the matching-group level along the T dimension is 0.645 – the average of b is 0.141 (with a standard deviation of 0.090) with $T = 6$ versus 0.190 (with a standard deviation of 0.137) for $T = 2$.[19] The difference is much smaller along the tax versus debt-financing dimension: an average of 0.187 (with a standard deviation of 0.097) in the former versus 0.192 (0.144) in the later and a corresponding p-value of the rank-sum test of 0.858.

As such, as discussed in Angeletos et al. (2021), this positive relationship between errors and revisions is rather generic and does not allow us to characterize which deviation(s) of the FIRE model is at play. Several distinct expectation models produce such a positive coefficient and we may rationalize it by assuming heterogeneous information within the context of otherwise RE agents or by considering a departure from RE, such as adaptive expectations. In the context of our experiment, the controlled lab environment ensures that information is common knowledge but variations in interpretation – whether due to heterogeneous cognitive abilities, heterogeneous beliefs about others in a context of strategic uncertainty or across subjects – may not be ruled out.

We therefore conduct a test for the presence of adaptive forms of expectations, which is a common finding in backward-looking designed LtFEs. As soon as revisions are included in the RHS of (3), the further inclusion of lagged output, or any other statistics, should be redundant. In the third column of Table 1 for each horizon we choose to include the second lag of output to test for the relevance of backward-looking expectations among the subjects. Note that we choose the second lag because the first lag of output is directly determined by the aggregate expectations formed in $t-1$. In other words, Y_{t-1} is an explicit function of $\bar{E}_{t-1}(Y_{t+\tau})$ per the DGP (1). A positive relationship between the two would then merely reflect the mechanical functioning of the model, rather than adaptive behavior in expectations.

The second lag of output as a control variable in Regression (3) is never significant, as clear from Table 4 and Panel A of Table 8.[20]This result speaks against

[19]However, one must be cautious in giving a final interpretation to this result: the relatively small number of groups, 8 in each side, makes it hard to disentangle between a true null effect and an absence of statistical significance due to a lack of power. This is an inherent limitation to group experiments.

[20]When adding the second lag of output for each session separately, none of the coefficients are significant, except for one session where the coefficient is weakly significant. This result speaks against systematic adaptive of extrapolative behaviors among the participants.

Dependent variable: individual *ex post* forecast errors

	$\tau = 1$			$\tau = 3$		
	OLS	FE	FE	OLS	FE	FE
$\hat{b}$	-0.140***	-0.140***	-0.140***			
(sd)	(0.019)	(0.006)	(0.006)	(0.101)	(0.389)	(0.334)
y_{t-2}			-0.018**			-0.014***
(sd)			(0.008)			(0.002)

Notes: see Table 4.

Table 5: Evidence of individual over-reaction in the experimental economies

a strong adaptive or backward-looking component in the expectation formation of the subjects. This results contrasts with previous findings in the LtFE literature. We conjecture that this is due to the forward-looking nature of the information set and the experimental task in our experiment. We should be careful in not over-interpreting this finding though. This is not a test for adaptive expectations *per se*. It is rather that a significant relationship between the ex-post forecast errors and the second lag of output would have been a strong indication of adaptive expectations.

4.2 Over-reaction to news at the individual level

Using now individual forecast data, the survey literature has established that individual errors are predictable by their own past revisions with a negative coefficient. That is, in the context of Model (3), the coefficient b would be negative when conducting the estimation at the individual level, using panel data of each of the 160 participants time series of *ex post* errors and forecast revisions over the $80 - T$ periods where forecasts are incentivized. Such a finding indicates overreaction to news and stands in contradiction to models such as noisy and dispersed information models that rely on RE at the individual level; see, e.g., Bordalo et al. (2020).

Table 5 reports these individual level estimates for our dataset for $\tau = 1$ and $\tau = 3$. As the results for the other horizons are similar, they are deferred to Table 8, Panel B in Appendix B.

We find significantly negative coefficients at all horizons, which supports overreaction at the individual level. Looking into cross-treatment differences, a rank-sum test at the matching-group level reveals that the deviation from the FIRE model is more pronounced with $T = 2$ (where the average coefficient b is -0.188 with a standard deviation of 0.267) than with $T = 6$ (average of -0.075 with a standard deviation of 0.183), the p-value being < 0.01. There is no significant difference along the tax versus debt-financed dimension (the p-value is 0.807). These results are consistent with Finding 1 above.

Additionally, including the second lag of output yields contrasting results, where the coefficients is or is not significant, depending on the horizon considered. Yet, in any case, the magnitude of the estimated coefficient is small, which does not provide strong support for adaptive behavior.

From Sections 4.1 and 4.2, we obtain the second main finding:

Finding 2 (Experimental data and survey evidence) *Overall, the experimental data are consistent with the stylized facts from the survey literature. In particular:*

i) the data reflects some information rigidities and aggregate forecasts tend to under-react to news;

ii) individual forecasts tend to over-react to news, which stands in contrast with the assumption of individual rationality.

We shall finally note that, while our data are qualitatively consistent with the two aforementioned stylized facts from survey data, the estimated values of the coefficients associated to the forecast revisions are smaller (in absolute terms) than those arising from survey data. This difference in magnitude indicates that subjects face less from informational frictions than respondents of survey data. We could easily rationalize such difference. The experiment spans over a couple of hours only, over which dozens of repetitions of the forecasting tasks take place. The attention of the subjects is also monopolized by the forecasting task, while surveys span over months (or years) and attention may not be guaranteed to the same extent as in a controlled lab environment. Our claim is that the lab constitutes a best-case scenario. To the extent that real-world people may only pay less attention to information than our lab subjects, if deviations from the FIRE model are found in a lab, this is a strong indication that such deviations are systematic and wide-spread in naturally-occurring settings.

We now study in detail the expectation formation process of the subjects.

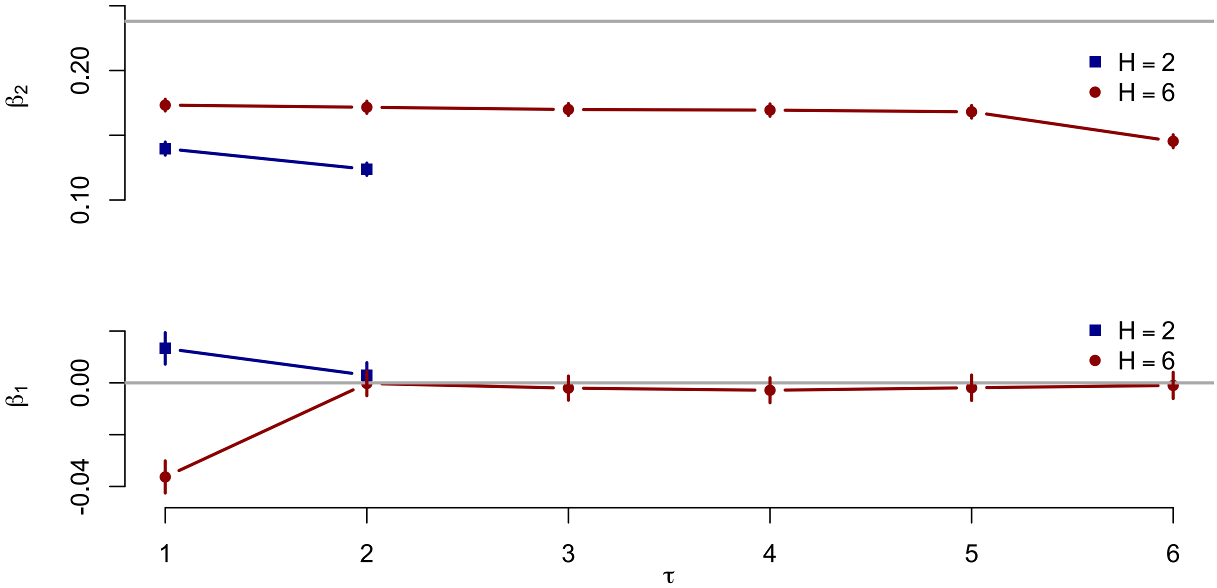
4.3 Is there evidence of extrapolative expectations?

We may use the various vintages of forecasts to compare how participants adjust their forecasts at various horizons when a news about a shock arrives. To do so, we adapt the method of Angeletos et al. (2021) that quantifies how forecasts react to structural disturbances to the context of our experiment where participants adjust their forecasts to announcements for future periods. This technique further investigates the presence of extrapolation behavior by measuring the potential persistence of otherwise one-time shocks due to the expectation channel.

We consider the following relation for each forecast horizon $\tau \in \{1, ..., T\}$ using a panel model across the 160 participants and the 80 periods of the experiment:

$$\hat{E}_{j,t}(Y_{t+\tau}) = \alpha_\tau + \beta_1 G_t + \beta_2 G_{t+\tau} + \gamma W_{i,t} + \epsilon_{t+\tau} \tag{4}$$

where we include in the vector of control variables W the lagged output Y_{t-1}, the lagged forecast of this output $E_{j,t-\tau-1}(Y_{t-1})$, a dummy for the tax versus debt-



Notes: The vertical bars (barely visible as the numbers are small) indicate two standard deviations around the estimated coefficients.

Figure 4: Forecast reactions to news at various horizons

financed spending treatment, a dummy for $T = 2$ versus $T = 6$ and a dummy for the lab location. We also include the announced (relevant) spending shock $g_{t+\tau}$[21] Because shocks have no persistence in the experiment, the estimated β_1 coefficients are zero under PF, no matter the horizon τ and the β_2 PF value is 0.243 under our calibration.

Figure 4 reports the results of these regressions. We can see two results. First, by looking into the estimates of β_1 (bottom panel), we find limited support for extrapolation behavior. These estimates are small in absolute terms and are only significantly different from zero for one-step-ahead forecasts. This means that at the time of the shock, the forecast for the period immediately following the shock is influenced by the shock, which should not be the case under PF. This implies some extrapolation but this extrapolation behavior does not span beyond the one-step-ahead forecasts. In other terms, subjects expect the shocks to quickly die out.

Second, as for the integration of news τ-period-ahead, we may look into the top panel of Figure 4. It is clear that there is a significant under-reaction, of comparable magnitude at all horizons, and this underreaction to the news shocks is greater for $T = 2$ than $T = 6$. Looking into $T = 6$, we also see that the reaction to news is lower for the six-period-ahead forecasts, which correspond to the forecasts submitted when subjects see the upcoming shocks for the first time, than for the five-period-ahead forecasts, that are submitted one period later.

Digging further into the data, there are in fact three times as many more zero output predictions for the six-period-ahead forecasts as for the five-period-ahead

[21]Knowing that shocks are sufficiently away from each other, there is no period such that non-zero shock in t and $t + h$ coexist in the information set of the participants.

forecasts, when accounting only for the forecasts that concern a period with a non-zero fiscal shock (16.4% versus 5.9%, with the test of the equality of proportions yielding a p-value < 0.001. This means that one out of six subjects is inattentive and misses the initial announcement of the fiscal shock, while only 5% remain so in the following period.

Figure 10 in Appendix further reports the estimates of (4) by separating the first 40 periods from the last 40 periods of the experiment. It is clear that considering the first half versus the second half of the experiment shows that the gap between the reaction of the participants to news and the PF value shrinks over time, which again brings another evidence of learning.

Finding 3 (Weak evidence of adaptive or extrapolative expectations) *The data does not provide strong support for adaptive or extrapolative expectations, contrary to previous LtFEs.*

Based upon Finding 3, we develop a Bayesian learning model to rationalize the behaviors observed in the experiment.

5 A Bayesian model of the participants' forecasts

The above discussion of our results suggests that participants steadily learn how to quantify the magnitude of the impact of fiscal shocks on output. In what follows, we develop a simple model of Bayesian updating to rationalize how subjects revise their beliefs about the influence of the state variables (future fiscal shocks) on output.

The model We formulate the problem faced by the forecasters in the experiment as follows. They observe government spending and try to estimate the coefficient pertaining to its effect on output based on a prior belief and new observations. From the instructions, subjects know that output may deviate from zero due to, either, movements in the average output expectations of the group or fiscal shocks.

Specifically, their model of the economy looks like:

$$Y_t = \theta_{i,t} G_t + e_{i,t}, \tag{5}$$

with $\theta_{i,t}$ the variable referring to the subject i's belief in time t and $e_{i,t}$ a noise term. Note that the linear form of Eq. (5) is consistent with the linear DGP of the experiment. We assume that e follows$\sim N(0, \sigma^2_{i,t})$ where $\sigma^2_{i,t}$ is the perceived variance by subject i in output that is due to the deviations of the expectations of the group from PF at time t.[22]

[22] An idiosyncratic noise term possibly reflects heterogeneous perception or limited attention.

Rewriting Eq. (5) gives:

$$\frac{Y_t}{G_t} = \theta_{i,t} + \frac{e_{i,t}}{G_t}, \tag{6}$$

noting that this model provides a model for the evolution of the beliefs of the subjects in the presence of fiscal shocks, namely for $G_t \neq 0$.

Now define $z_t = \frac{Y_t}{G_t}$ and $v_{i,t} = \frac{e_{i,t}}{G_t}$ so that we can write

$$z_t = \theta_{i,t} + v_{i,t} \tag{7}$$

This is a signal extraction problem without process noise that can be solved with a simplified version of the Kalman filter.

We may not observe the true beliefs of the subjects but the belief $\theta_{i,t}$ is the available proxy for the subject's true, unobservable belief about the effect of the shocks on output. To simplify the exposition, in the sequel, we refer to $\theta_{i,t}$ as the *individual proxied belief*. Hence, let $\hat{\theta}_{i,t}$ be the current (prior) estimate of the subject's belief θ and $\hat{P}_{i,t}$ the subject's uncertainty about their estimate. Noting that $e_{i,t} \sim N(0, \frac{\sigma^2_{i,t}}{G^2_t})$, the Kalman gain reduces to:

$$K_{i,t} = \frac{\hat{P}_{i,t}}{\hat{P}_{i,t} + \frac{\sigma^2_{i,t}}{G^2_t}}. \tag{8}$$

Moreover, $\hat{\theta}_{i,t}$ and $\hat{P}_{i,t}$ evolve as:

$$\hat{\theta}_{i,t+1} = \hat{\theta}_{i,t} + K_{i,t}(z_t - \hat{\theta}_{i,t}), \tag{9}$$

$$\hat{P}_{i,t+1} = (1 - K_{i,t})\hat{P}_{i,t}. \tag{10}$$

Estimation strategy We take this learning model to our experimental data as follows. For each of the 160 subjects, the initial belief of the subject, that we denote $\hat{\theta}_{i,0}$, is pinned down by their output forecast for the first period featuring a government spending shock, namely period 17. That is, $\hat{\theta}_{i,0}$ is set equal to their two-period-ahead forecast submitted in period 15 in the treatments with $T = 2$ and their six-period-ahead forecast submitted in period 11 when $T = 6$, divided by the magnitude of the first shock. The initial uncertainty, $\hat{P}_{i,0}$, determines the speed of updating. We estimate its value for each individual so as to best match the dynamic updating of their forecasts in the experimental data via maximum of likelihood.

The last parameter in the model is the perceived variance, denoted by $\sigma^2_{i,t}$, that arises due to strategic uncertainty in the otherwise deterministic environment. In other words, it is group-dependent and results from fluctuations in the average output expectations of the group that, in turn, generate deviations of output from its PF path. For this reason, we choose a group-specific rather than a subject-specific value and pin down $\sigma^2_{i,t}$ by using the experimental data as follows. We use the observed sample variance of output in the absence of any government spending shocks by computing, for each group, the variance of output during the first periods of each experimental economy up until the first announcement of a spending shock.[23] For each participant, we make use of the seven forecast data points, each corresponding to the change of output forecast following the first announcement of each of the seven shocks.[24]

Outcomes Figure 5a contrasts the distribution of the final proxied beliefs $\{\theta_{7,i}\}$ (namely the last two-period or six period ahead forecasts for the period of the last fiscal shock, divided by the value of this shock) with the distribution of the final estimated beliefs, that we denote by $\{\hat{\theta}_{7,i}\}$, obtained by maximum of likelihood using the model previously described. The matching group-level KS test gives a p-value of 0.699: there is no statistical difference between the distribution of the estimated and the observed beliefs, which indicates a good fit of the model. This is an interesting complement to Finding 3.

Figure 5b further compares the distribution of the 160 initial proxied beliefs $\{\theta_{0,i}\}$ and the distribution of the final proxied beliefs. The matching group-level KS test returns a p-value < 0.01 between the two distributions, which is a clear sign of learning. As also salient from looking at the two distributions, there is a shift of the subjects' proxied beliefs towards the PF equilibrium value. Yet, the one-sided rank-sum test of the difference between the distribution of the final proxied beliefs and the equilibrium value θ returns a p-value < 0.01, which indicates that subjects fall short of learning the full extent of the effect of the spending shocks on output.[25]

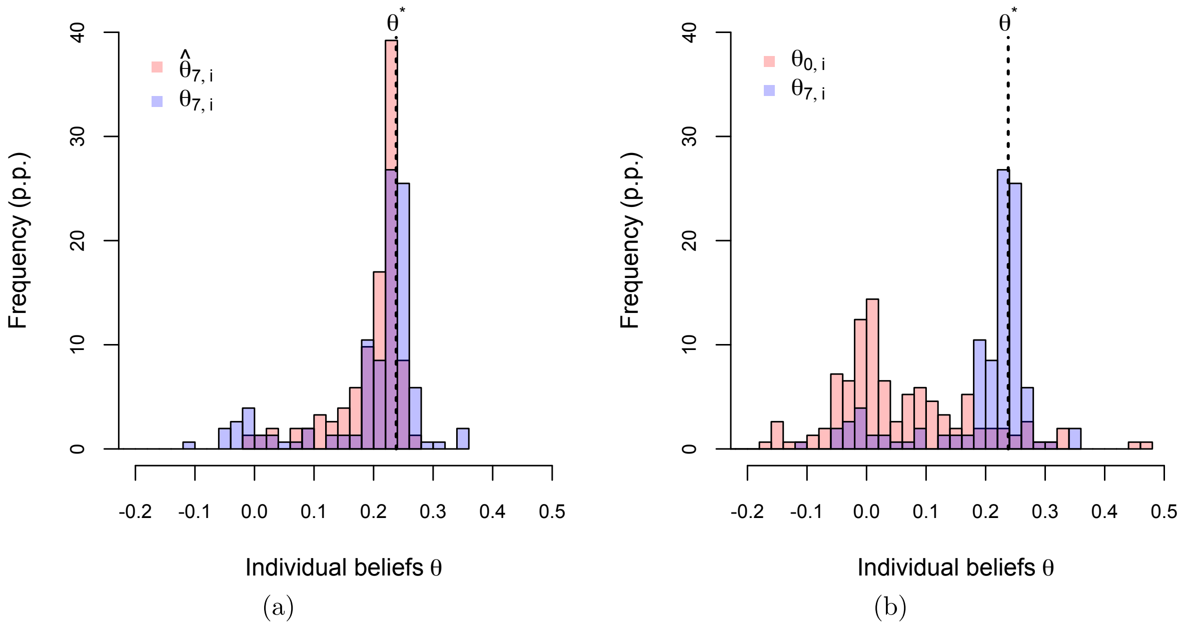
Table 6 studies the determinants of the fit and the speed of learning among the subjects using panel regression models. Columns (I) and (II) report the results when (minus) the log-likelihood of the estimation is used as a dependent variable. While such metrics does not have an explicit interpretation, it measures the goodness of fit, where higher values indicate a better fit of the Bayesian learning model to the forecast data for a particular subject. The regressors include the treatment variables (Column I), together with the individual characteristics of each subject (Column II).

We find two main insights. First, longer-horizon environments lead to a better fit than short-horizon ones, which means that subjects form forecasts that are closer to a Bayesian model in the $T = 6$-treatments than in the $T = 2$-treatments. This is again in line with Finding 1. By contrast, whether spending is tax- or debt-financed and the location of the subject pool does not influence the fit.

[23]This method supposes that each subject in a given group perceives the same amount of strategic uncertainty. The results that we discuss below are not sensitive to the alternative that consists in estimating the individual values of $\sigma^2_{i,t}$ so as to best match individual time series. For this reason, we rather present the result of the more parsimonious alternative.

[24]The results are robust to using all forecasts, not just the first forecast made after the announcement of the shock.

[25]The result is the same if we take the final estimates of the proxied beliefs $\hat{\theta}_{i,7}$, rather than the final proxied beliefs $\theta_{i,7}$.



Notes: It should be recalled that the true individual beliefs are not observable, hence we use proxied beliefs via the variable θ.

Figure 5: Distributions of individual proxied beliefs in the experiment

Second, some individual characteristics are significantly associated to a better fit of the Bayesian model. These are the previous experience of the subjects with LtFE, the level of effort that they declare having invested when making forecasts and the choice of ignoring irrelevant variables, such as the level of debt.[26]

Columns (III) and (IV) reproduce the same analysis with the average Kalman gain estimated for each subject as the dependent variable. This variables has a direct interpretation and measures how fast a subject revises their belief, where higher values indicate a faster learning process.[27] Again, longer horizons favor faster learning. Subjects who tend to ignore the fiscal shocks or react to them with a delay end up learning more slowly than subjects who declare having quickly reacted to the shocks. Subjects in Bamberg and subjects with less experience also tend to learn faster. This may be explained by the fact that subjects in Amsterdam participated in more experiments overall, and in particular more LtFEs than subjects in Bamberg. A good fit needs not be incompatible with a lower gain value: it may be that some subjects are well-described by a Bayesian updating mechanism but simply update their beliefs more slowly than others. It may be the case also if their initial belief is already quite accurate.

[26]We recall that the level of debt or tax does not influence the impact of fiscal shocks on output in the underlying model of the economy.

[27]Results are robust to using the final gain value instead of the average over the seven shocks.

	Dependent variable:			
	- log-likelihood		average gain $\{K_{i,t}\}_{t=1,\ldots,7}$	
	(I)	(II)	(III)	(IV)
Dummy for $T = 6$ (1)	1.883***	2.193***	0.044**	0.054***
	(0.582)	(0.551)	(0.021)	(0.017)
Dummy for debt-financed (1)	0.191	−0.031	−0.011	−0.008
	(0.561)	(0.527)	(0.023)	(0.018)
Dummy for Bamberg (1)	−0.741	−1.283	0.042*	0.043**
	(0.735)	(0.860)	(0.022)	(0.020)
Q1 (effort level)		1.505***		−0.006
		(0.462)		(0.008)
Q2 (immediate reaction to spending shocks)		0.591		0.008
		(0.681)		(0.011)
Q3 (delayed reaction to spending shocks)		−0.045		−0.033***
		(0.364)		(0.008)
Q4 (ignoring spending shocks)		−0.852		−0.041***
		(0.700)		(0.010)
Q5 (ignoring government debt)		0.745**		0.016**
		(0.331)		(0.007)
Q6 (ignoring taxes)		0.050		0.004
		(0.299)		(0.008)
Q7 (high cognitive load confusion		−0.322		−0.027***
		(0.583)		(0.008)
Dummy for no experience		−1.831*		0.044**
		(1.025)		(0.020)
Time to complete the quiz (cognitive ability)		0.0004		−0.00001
		(0.001)		(0.00002)
Constant	10.435***	1.260	0.312***	0.422***
	(2.645)	(5.237)	(0.064)	(0.077)
Demographics	Yes	Yes	Yes	Yes
# of observations	156	156	156	156
R^2	0.073	0.216	0.130	0.427

Notes: *p<0.1; **p<0.05; ***p<0.01. Standard errors are reported below the coefficients in parentheses. Errors are clustered at the group level. Forecasts submitted during the last T periods are non-incentivized, hence discarded. Demographics corresponds to the age, the field of study, the geographical origin and the gender. None of these variables are significant (the lowest p-value is 0.105). The exact questions are given in Appendix C. Self-reported answers range from 1 (do not agree at all) to 5 (fully agree).

Table 6: OLS models of the Bayesian updating mechanism

6 Conclusion

We set up a LtFE that adds two dimensions to this class of experiments. First, subjects forecast every future realization over an entire horizon, i.e. beyond the usually elicited one-period-ahead prediction, such that the revision of several vintages is a large part of their forecasting task. Second, we project information about the future path of the relevant shocks within this forecasting horizon. These shocks are treated as news and we study the reaction of lab forecasts to these news.

We find that these two elements together induce somewhat forward-looking behavior of the subjects, in the sense that we find little support for simple and fully backward-looking models of expectations, such as adaptive or extrapolative expectations. This result stands in stark contrast with most of the related experimental literature that has been concerned with studying how past developments are reflected into short-run forecasts.

We further complement the recent efforts to strengthen the external validity of the lab environment by comparing survey and lab forecast data. We find that the lab data are qualitatively consistent with the two main stylized facts from the survey literature, namely information rigidities and underreaction to news at the aggregate level and overreaction to news in the individual forecasts, which speaks against individual rationality. The order of magnitude of the effects are nevertheless smaller than in forecast data, which may be explained by the somewhat artificial lab setting where attention is monopolized by the forecasting task.

Overall, we find that subjects display clear learning dynamics. Moreover, the longer the planning horizon, the stronger the learning dynamics and the smaller the forecast errors. We rationalize the experimental data within a Bayesian updating model where agents successively revise their beliefs about the effect of spending shocks on output and converge asymptotically towards the perfect foresight equilibrium value of the model. We find that this model provides a better description of the behaviors in long-horizon than in short-horizon environments and among attentive and experienced subjects rather than novice or confused subjects.

Of course, one may acknowledge that our setting is particularly stylized. The structure of the shocks and their effects on the forecasted variable is much simplified compared to complex economic models, or even the real-world economies. Nevertheless, our experiment suggests that if we project information about the (near) future that is in line with the horizon people are interested in, adaptive behavior dampens. People may then learn to incorporate news in a way that is closer to the FIRE benchmark than in purely backward-looking settings. A longer horizon seems to facilitate this learning process. This begs the question of which horizon is relevant for people in the real world and whether policy, via making announcements over the chosen horizon for instance, may influence this horizon to shape expectations and, indirectly, economic outcomes.

References

Hassan Afrouzi, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar. Overreaction in expectations: Evidence and theory. *Available at SSRN*, 2021. 5

Steffen Ahrens, Joep Lustenhouwer, and Michele Tettamanzi. The Stabilizing Effects of Publishing Strategic Central Bank Projections. *Macroeconomic Dynamics*, 2022. doi: 10.1017/S1365100521000687. 5

George-Marios Angeletos, Zhen Huo, and Karthik A Sastry. Imperfect macroeconomic expectations: Evidence and theory. *NBER Macroeconomics Annual*, 35 (1):1–86, 2021. 4, 21, 23

Mikhail Anufriev, Aleksei Chernulich, and Jan Tuinstra. Asset Price Volatility and Investment Horizons: An Experimental Investigation. Working Papers 20200053, New York University Abu Dhabi, Department of Social Science, August 2020. 5

Olivier Armantier, Scott Nelson, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar. The price is right: Updating inflation expectations in a randomized price information experiment. *Review of Economics and Statistics*, 98(3):503–523, 2016. 4

Te Bao, Cars Hommes, and Jiaoying Pei. Expectation formation in finance and macroeconomics: A review of new experimental evidence. *Journal of Behavioral and Experimental Finance*, 32:100591, 2021. 5

Robert B Barsky and Eric R Sims. News shocks and business cycles. *Journal of monetary Economics*, 58(3):273–289, 2011. 4

Carola Binder and Alex Rodrigue. Household informedness and long-run inflation expectations: Experimental evidence. *Southern Economic Journal*, 85(2):580–598, 2018. 4

Pedro Bordalo, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–82, 2020. 4, 22

Alberto Cavallo, Guillermo Cruces, and Ricardo Perez-Truglia. Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics*, 9(3):1–35, 2017. 4

Daniel L Chen, Martin Schonger, and Chris Wickens. otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016. 13

Olivier Coibion and Yuriy Gorodnichenko. What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116–159, 2012. 4

Olivier Coibion and Yuriy Gorodnichenko. Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78, 2015. 4, 19

Olivier Coibion, Yuriy Gorodnichenko, and Saten Kumar. How do firms form their expectations? new survey evidence. *American Economic Review*, 108(9):2671–2713, 2018. 4

Olivier Coibion, Yuriy Gorodnichenko, Edward S Knotek II, and Raphael Schoenle. Average inflation targeting and household expectations. Technical report, National Bureau of Economic Research, 2020. 5

Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 2022. 5

Camille Cornand and Paul Hubert. On the external validity of experimental inflation forecasts: A comparison with five categories of field expectations. *Journal of Economic Dynamics and Control*, 110:1037–46, 2020. 3

G. Evans, C. Hommes, B. McGough, and I. Salle. Are Long-Horizon Expectations (De-)Stabilizing? Theory and Experiments. Staff Working Papers 19-27, Bank of Canada, August 2019. 3, 5, 7, 16

George W Evans and Seppo Honkapohja. *Learning and expectations in macroeconomics*. Princeton University Press, 2012. 6

Andreas Fuster and Basit Zafar. Survey experiments on economic expectations. Technical report, National Bureau of Economic Research, 2022. 4

Ingar Haaland, Christopher Roth, and Johannes Wohlfart. Designing information provision experiments. *Journal of Economic Literature*, 2022. 4

Ernan Haruvy, Yaron Lahav, and Charles N Noussair. Traders' expectations in asset markets: experimental evidence. *American Economic Review*, 97(5):1901–1920, 2007. 5

Oleksiy Kryvtsov and Luba Petersen. Expectations and monetary policy: Experimental evidence. *Unpublished working paper, Simon Fraser University*, 2015. 5

Oleksiy Kryvtsov and Luba Petersen. Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics*, 117:760–780, 2021. 5

Joep Lustenhouwer. Fiscal stimulus in expectations-driven liquidity traps. *Journal of Economic Behavior & Organization*, 177:661–687, 2020. 6

Joep Lustenhouwer and Kostas Mavromatis. The effects of fiscal policy when planning horizons are finite. *De Nederlandsche Bank Working Paper*, 2021. 6

Philipp Marquardt, Charles N Noussair, and Martin Weber. Rational expectations in an experimental asset market with shocks to market trends. *European Economic Review*, 114:116–140, 2019. 5

Thomas Meissner and Davud Rostam-Afschar. Learning ricardian equivalence. *Journal of Economic Dynamics and Control*, 82:273–288, 2017. 6

F. Mokhtarzadeh and L Petersen. Coordinating expectations through central bank projections. *Experimental Economics*, 24:883–918, 2021. 5

Ryan Rholes and Luba Petersen. Should central banks communicate uncertainty in their projections? *Journal of Economic Behavior & Organization*, 183:320–341, 2021. 5

Isabelle Salle. Comment on “No firm is an island? How industry conditions shape firms’ expectations” by Philippe Andrade, Olivier Coibion, Erwan Gautier and Yuriy Gorodnichenko. *Journal of Monetary Economics*, 125:57–61, 2022. 2

Michael Woodford. Monetary policy analysis when planning horizons are finite. *NBER Macroeconomics Annual*, 33(1):1–50, 2019. 6, 34, 36

A Derivations of the underlying model

This section derives the DGP of the underlying experimental economy, namely Equation (1) in the main text.

A.1 Households

Households maximize their discounted utility of consumption and leisure over their planning horizon T. Households also need to value their individual state they expect to end up in at the end of these T periods. The finite-horizon setting that we use here assumes bounded rationality in so far as households are assumed not to be able to rationally induce (by solving the model forward until infinity) their state value in period $T+1$. Instead, we assume that they use a rule of thumb to evaluate the value of their state. This rule of thumb is assumed to be correctly specified in so far as households only take into account their bond holdings to evaluate their future state because they are aware that this is the only relevant state variable. Their maximization problem becomes:

$$\max_{\{C_s^i, H_s^i, B_{s+1}^i\}_{s=t}^{t+T}} \tilde{E}_t^i \sum_{s=t}^{t+T} \beta^{s-t} u(C_s^i, H_s^i) + \beta^{T+1} V(\frac{B_{t+T+1}^i}{P_{t+T}}), \tag{11}$$

subject to

$$P_s C_s^i + \frac{B_{s+1}^i}{1+i_s} \leq W_s H_s^i + B_s^i + P_s \Xi_s - P_s \tau_s, \qquad s = t, t+1, ...t+T, \tag{12}$$

where C_t^i, H_t^i and B_{t+1}^i are respectively the consumption, labor and nominal bond holdings from household i decided upon in period t. Ξ_t are real profits from firms, which are assumed to be equally distributed among households and W_t is the nominal wage rate. Finally, P_t is the price level in period t, and τ_t are lump-sum taxes.

We assume that households have CRRA preferences with relative risk aversion σ and Frisch elasticity of labor supply η.[28] The functional form of the value function $V(.)$ is then given by:

$$V(x) = \frac{1}{1-\beta} \left[\frac{(\Lambda + (1-\beta)x)^{1-\sigma}}{1-\sigma} + \frac{(\bar{H})^{1+\eta}}{1+\eta} \right], \tag{13}$$

where $\Lambda = \bar{w}\bar{H} + \bar{\Xi} - \bar{\tau}$ is the households' steady-state net income.

The assumed value function is the continuation value that solves the Bellman equation:

$$V(x) = \max_c \{U(C, \bar{H}) + \beta V(x')\}, \quad s.t. \quad x' = \frac{1}{\beta} \left[\bar{w}\bar{H} + x + \bar{\Xi} - \bar{\tau} - C \right] \tag{14}$$

As explained in Woodford (2019), this problem describes the optimal intertem-

[28] Our utility function hence takes the functional form $u(C_s^i, H_s^i) = \frac{(C_s^i)^{1-\sigma}}{1-\sigma} - \frac{(H_s^i)^{1+\eta}}{1+\eta}$.

poral consumption decision of the household assuming that taxes, wages, hours worked and profits all equal their steady state values. Households then make optimal decisions at the steady state. This implies that households understand the role of bonds in consumption smoothing over their horizons. For instance, if they expect to have more bonds at the end of the horizon, they realize how this will allow them to consume more after their horizon and hence obtain more utility. However, this model of bounded rationality considers that households are not sophisticated enough to plan how their hours worked, wages and profits would change after their horizon if they would have different consumption levels. The value function hence captures partly how future utility flows depend on the end-of-horizon wealth, but in a boundedly rational manner that only approximates the true value function.

It is convenient to divide the budget constraint by P_s:

$$C_s^i + \frac{B_{s+1}^i}{(1+i_s)P_s} \leq w_s H_s^i + \frac{B_s^i}{P_s} + \Xi_s - \tau_s, \qquad s = t, t+1, ...t+T. \tag{15}$$

We obtain the following first-order conditions of the maximization problem:

$$(H_s^i)^\eta = (C_s^i)^{-\sigma} w_s, \qquad s = t, t+1, ...t+T, \tag{16}$$

$$(C_s^i)^{-\sigma} = \beta \frac{(1+i_s)(C_{s+1}^i)^{-\sigma}}{\Pi_{s+1}}, \qquad s = t, t+1, ...t+T-1, \tag{17}$$

$$(C_{t+T}^i)^{-\sigma} = \beta(1+i_{t+T})(\Lambda + (1-\beta)\frac{B_{t+T+1}^i}{P_{t+T}})^{-\sigma}. \tag{18}$$

Next, it is convenient to define a measure of real bond holdings scaled by steady-state output: $b_t = \frac{B_t}{P_{t-1}\bar{Y}}$. Substituting for this expression in (15) and (18) yields:

$$C_s^i + \bar{Y}\frac{b_{s+1}^i}{1+i_s} \leq w_s H_s^i + \frac{\bar{Y} b_s^i}{\Pi_s} + \Xi_s - \tau_s, \qquad s = t, t+1, ...t+T, \tag{19}$$

and

$$(C_{t+T}^i)^{-\sigma} = \beta(1+i_{t+T})(\Lambda + (1-\beta)\bar{Y} b_{t+T+1}^i)^{-\sigma}, \tag{20}$$

where $\Pi_s = P_s/P_{s-1}$ is gross inflation in any period s.

A.2 Firms

The model is populated by a continuum of monopolistically competitive firms producing the final differentiated goods. Each firm has a linear technology with labor as its only input:

$$Y_t(j) = AH_t(j), \tag{21}$$

where A represents aggregate productivity, which is assumed to be constant.

Firms are run by households, and hence operate under the same finite-planning horizons. Consequently, they form expectations about their marginal costs and the demand for their product for T-period-ahead. We assume price stickiness as follows. In each period, only a fraction $(1-\omega)$ of firms can change their price. The problem of firm j that can reset its price is then to maximize the discounted value of its nominal profits for the next T periods:

$$\tilde{E}_t^j \left(\sum_{s=0}^{T} \omega^s Q_{t,t+s}^j \left[p_t(j) Y_{t+s}(j) - P_{t+s} mc_{t+s} Y_{t+s}(j) \right] + \omega^{T+1} \beta^{T+1} \left(\frac{\bar{C}}{C_t^j} \right)^{-\sigma} P_t \tilde{V}\left(\frac{p_t(j)}{P_{t+T}}\right) \right), \tag{22}$$

where

$$Q_{t,t+s}^j = \beta^s \left(\frac{C_{t+s}^j}{C_t^j} \right)^{-\sigma} \frac{P_t}{P_{t+s}}, \tag{23}$$

is the stochastic discount factor of the household that runs firm j.

Following Woodford (2019), the functional form of the value function $\tilde{V}(r)$ reads as:

$$\tilde{V}(r) = \frac{1}{1-\omega\beta} \left(r^{1-\theta} \bar{Y} - r^{-\theta} \bar{Y} \bar{w} \right). \tag{24}$$

Using the demand for good j, the firm's profit maximization problem writes as follows:

$$\begin{aligned} \max \tilde{E}_t^j \Bigg(& \sum_{s=0}^{T} \omega^s \beta^s \left(\frac{C_{t+s}^j}{C_t^j} \right)^{-\sigma} P_t \left[\left(\frac{p_t(j)}{P_{t+s}} \right)^{1-\theta} Y_{t+s} - mc_{t+s} \left(\frac{p_t(j)}{P_{t+s}} \right)^{-\theta} Y_{t+s} \right] \\ & + \frac{(\omega\beta)^{T+1}}{1-\omega\beta} \left(\frac{\bar{C}}{C_t^j} \right)^{-\sigma} P_t \left[\left(\frac{p_t(j)}{P_{t+T}} \right)^{1-\theta} \bar{Y} - \bar{mc} \left(\frac{p_t(j)}{P_{t+T}} \right)^{-\theta} \bar{Y} \right] \Bigg). \end{aligned} \tag{25}$$

The first order condition for $p_t(j)$ is

$$\begin{aligned} & \tilde{E}_t^j \sum_{s=0}^{T} \omega^s \beta^s \left(\frac{C_{t+s}^j}{C_t^j} \right)^{-\sigma} \frac{P_t}{P_{t+s}} Y_{t+s} \left[(1-\theta) \left(\frac{p_t^*(j)}{P_{t+s}} \right)^{-\theta} + \theta mc_{t+s} \left(\frac{p_t^*(j)}{P_{t+s}} \right)^{-1-\theta} \right] \\ & + \frac{(\omega\beta)^{T+1}}{1-\omega\beta} \left(\frac{\bar{C}}{C_t^j} \right)^{-\sigma} \frac{P_t}{P_{t+T}} \bar{Y} \left[(1-\theta) \left(\frac{p_t^*(j)}{P_{t+T}} \right)^{-\theta} + \theta \bar{mc} \left(\frac{p_t^*(j)}{P_{t+T}} \right)^{-1-\theta} \right] = 0, \end{aligned} \tag{26}$$

where $p_t^*(j)$ is the optimal price for firm j if it can re-optimize in period t.

Multiplying by $\frac{(C_t^j)^{-\sigma}}{P_t} \frac{p_t^*(j)^{1+\theta}}{1-\theta}$ gives:

$$\begin{aligned} & \tilde{E}_t^j \sum_{s=0}^{T} \omega^s \beta^s \left(C_{t+s}^j \right)^{-\sigma} \frac{Y_{t+s}}{P_{t+s}} \left[p_t^*(j) P_{t+s}^{\theta} - \frac{\theta}{\theta-1} mc_{t+s} P_{t+s}^{1+\theta} \right] \\ & + \frac{(\omega\beta)^{T+1}}{1-\omega\beta} (\bar{C})^{-\sigma} \frac{\bar{Y}}{P_{t+T}} \left[p_t^*(j) P_{t+T}^{\theta} - \frac{\theta}{\theta-1} \bar{mc} P_{t+T}^{1+\theta} \right] = 0, \end{aligned} \tag{27}$$

which can be written as:

$$\frac{p_t^*(j)}{P_t}\left[\tilde{E}_t^j\sum_{s=0}^{T}\omega^s\beta^s\left(C_{t+s}^j\right)^{-\sigma}\left(\frac{P_{t+s}}{P_t}\right)^{\theta-1}Y_{t+s}+\frac{(\omega\beta)^{T+1}}{1-\omega\beta}\bar{Y}(\bar{C})^{-\sigma}\left(\frac{P_{t+T}}{P_t}\right)^{\theta-1}\right] \tag{28}$$

$$=\frac{\theta}{\theta-1}\left[\tilde{E}_t^j\sum_{s=0}^{T}\omega^s\beta^s\left(C_{t+s}^j\right)^{-\sigma}\left(\frac{P_{t+s}}{P_t}\right)^{\theta}Y_{t+s}mc_{t+s}+\frac{(\omega\beta)^{T+1}}{1-\omega\beta}\bar{Y}(\bar{C})^{-\sigma}\bar{mc}\left(\frac{P_{t+T}}{P_t}\right)^{\theta}\right].$$

Finally, the aggregate price level evolves as:

$$P_t=[\omega P_{t-1}^{1-\theta}+(1-\omega)\int_0^1 p_t^*(j)^{1-\theta}dj]^{\frac{1}{1-\theta}}. \tag{29}$$

A.3 Government and market clearing

The government issues bonds B and levies lump-sum taxes (τ_t) to finance its spending (G_t). Its budget constraint is given by:

$$\frac{B_{t+1}}{1+i_t}=P_tG_t-P_t\tau_t+B_t, \tag{30}$$

where $H_t=\int H_t^i di$ and $B_t=\int B_t^i di$ correspond, respectively, to aggregate labor and aggregate bond holdings.

Dividing both sides by $P_t\bar{Y}$ gives:

$$\frac{b_{t+1}}{1+i_t}=\frac{G_t}{\bar{Y}}-\frac{\tau_t}{\bar{Y}}+\frac{b_t}{\Pi_t}, \tag{31}$$

where $b_t=\frac{B_t}{P_{t-1}\bar{Y}}$ is the ratios of debt-to-steady-state output.

Market clearing is given by:

$$Y_t=C_t+G_t\equiv C_t+\bar{Y}g_t. \tag{32}$$

Monetary policy is defined by a Taylor-type interest rule where the authorities respond to current inflation:

$$\frac{1+i_t}{1+\bar{i}}=\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\phi_1}. \tag{33}$$

A.4 Computation of the steady state

In this section, we derive the steady state around which the model is log-linearized, where gross inflation equals 1 and debt equals 0.

Evaluating (28) at the zero-inflation steady state gives:

$$\bar{mc}=\frac{\theta-1}{\theta}. \tag{34}$$

From the first-order conditions of the households, it follows that, in this steady state, we must have:

$$1+\bar{i}=\frac{1}{\beta}. \tag{35}$$

Furthermore, it follows from (21) that:

$$\bar{H}A=\bar{Y}. \tag{36}$$

Next, we solve the steady-state aggregate-resource constraint (32) for consumption, and write:

$$\bar{C}=\bar{Y}-\bar{G}=\bar{Y}(1-\bar{g}). \tag{37}$$

Plugging these steady-state labor and consumption levels in the steady-state version of (16) gives:

$$\bar{w}=\frac{\bar{Y}^{\eta}(\bar{Y}(1-\bar{g}))^{\sigma}}{A^{\eta}}=\frac{\bar{Y}^{\eta+\sigma}(1-\bar{g})^{\sigma}}{A^{\eta}}=\frac{\theta-1}{\theta}, \tag{38}$$

where the last equality follows from $\bar{mc}=\bar{w}$ and (34).

We can thus write:

$$\bar{Y}=(\frac{\theta-1}{\theta}\frac{A^{\eta}}{(1-\bar{g})^{\sigma}})^{\frac{1}{\eta+\sigma}}. \tag{39}$$

Next, we turn to the government-budget constraint. At the steady state, (31) reduces to:

$$0=\bar{b}=\frac{\bar{\tau}-\bar{G}}{\bar{Y}(1-\beta)}, \tag{40}$$

where we use (35) to substitute for the interest rate. Steady-state taxes are therefore pinned down by the steady-state government spending level as:

$$\bar{\tau}=\bar{G}=\bar{g}\bar{Y}. \tag{41}$$

We now log-linearize the model around this steady state.

A.5 Log-linear model

A.5.1 Optimal consumption decision

The log-linearized optimality conditions (including the budget constraints) are given by:

$$\hat{C}^i_s=\hat{C}^i_{s+1}-\frac{1}{\sigma}(E^i_t i_s-E^i_t\pi_{s+1}), \qquad s=t,t+1,...t+T-1, \tag{42}$$

$$\tilde{b}^i_{t+T+1}=\frac{1-\bar{g}}{1-\beta}\hat{C}^i_{t+T}+\frac{1-\bar{g}}{(1-\beta)\sigma}E^i_t i_{t+T}, \tag{43}$$

$$\eta\hat{H}^i_s=-\sigma\hat{C}^i_s+E^i_t\hat{w}_s, \qquad s=t,t+1,...t+T, \tag{44}$$

$$\tilde{b}^i_{s+1} = \frac{\bar{w}}{\beta}(E^i_t\hat{w}_s + \hat{H}^i_s) + \frac{1}{\beta}\tilde{b}^i_s + \frac{\bar{\Xi}}{\bar{Y}\beta}E^i_t\hat{\Xi}_s - \frac{\bar{\tau}}{\bar{Y}\beta}E^i_t\hat{\tau}_s - \frac{1-\bar{g}}{\beta}\hat{C}^i_s, \tag{45}$$
$$s = t, t+1, ...t+T,$$

where we use $\bar{H} = \bar{Y}$ and $\frac{\bar{C}}{\bar{Y}} = 1 - \bar{g}$.

Iterating the log-linearized budget constraints from period $t+T$ backward and multiplying both sides by β^{T+1} gives:

$$\beta^{T+1}\tilde{b}^i_{t+T+1} = \tilde{b}^i_t - (1-\bar{g})\sum_{s=0}^{T}\beta^s(\hat{C}^i_{t+s}) + \frac{\bar{\Xi}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E^i_t\hat{\Xi}_{t+s}) - \frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E^i_t\hat{\tau}_{t+s})$$
$$+ \bar{w}\sum_{s=0}^{T}\beta^s(E^i_t\hat{w}_{t+s} + \hat{H}^i_{t+T-s}).$$

We then plug in $\tilde{b}^i_{t+T+1}$ from (43) and labor from (44) to obtain:

$$\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}\hat{C}^i_{t+T} + \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}E^i_t i_{t+T} = \tilde{b}_t - (1-\bar{g})\sum_{s=0}^{T}\beta^s(\hat{C}^i_{t+s}) + \frac{\bar{\Xi}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E^i_t\hat{\Xi}_{t+s})$$
$$- \frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E^i_t\hat{\tau}_{t+s}) + \bar{w}\sum_{s=0}^{T}\beta^s(E^i_t\hat{w}_{t+s} - \frac{\sigma}{\eta}\hat{C}^i_{t+s} + \frac{1}{\eta}E^i_t\hat{w}_{t+s}). \tag{46}$$

Next, we use the Euler equation to substitute for future consumption. Iterating the Euler equation gives:

$$\hat{C}^i_{t+s} = \hat{C}^i_t + \sum_{j=0}^{s-1}\frac{1}{\sigma}(E^i_t i_{t+j} - E^i_t\pi_{t+j+1}), \qquad T - s \geq 1. \tag{47}$$

Rearranging (46) and substituting for future consumption gives:

$$\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}\left(\hat{C}^i_t + \sum_{j=0}^{T-1}\frac{1}{\sigma}(E^i_t i_{t+j} - E^i_t\pi_{t+j+1})\right) + \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}E^i_t i_{t+T} \tag{48}$$
$$=\tilde{b}^i_t - \left(\frac{\sigma}{\eta}\bar{w} + (1-\bar{g})\right)\sum_{s=0}^{T}\beta^s\left(\hat{C}^i_t + \sum_{j=0}^{s-1}\frac{1}{\sigma}(E^i_t i_{t+j} - E^i_t\pi_{t+j+1})\right)$$
$$+ \frac{\bar{\Xi}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E^i_t\hat{\Xi}_{t+s}) - \frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E^i_t\hat{\tau}_{t+s}) + \bar{w}\sum_{s=0}^{T}\beta^s((1+\frac{1}{\eta})(E^i_t\hat{w}_{t+s}).$$

Taking contemporaneous consumption to one side of the equation gives the current decision of consumer i:

$$\left(\frac{\sigma}{\eta}\bar{w}\frac{1-\beta^{T+1}}{1-\beta}+\frac{1-\bar{g}}{(1-\beta)}\right)\hat{C}_t^i = \tag{49}$$

$$\tilde{b}_t^i+\bar{w}\sum_{s=0}^{T}\beta^s((1+\frac{1}{\eta})(E_t^i\hat{w}_{t+s})+\frac{\bar{\Xi}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E_t^i\hat{\Xi}_{t+s})-\frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(E_t^i\hat{\tau}_{t+s})$$

$$-\left(\frac{\sigma}{\eta}\bar{w}+(1-\bar{g})\right)\sum_{s=1}^{T}\beta^s\sum_{j=0}^{s-1}\frac{1}{\sigma}(E_t^i i_{t+j}-E_t^i\pi_{t+j+1})$$

$$-\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}\frac{1}{\sigma}\sum_{j=0}^{T-1}(E_t^i i_{t+j}-E_t^i\pi_{t+j+1})-\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}E_t^i i_{t+T}.$$

Aggregating this equation over all households yields an expression for aggregate consumption as a function of aggregate expectations about aggregate variables only, as follows:

$$\left(\frac{\sigma}{\eta}\bar{w}\frac{1-\beta^{T+1}}{1-\beta}+\frac{1-\bar{g}}{(1-\beta)}\right)\hat{C}_t = \tag{50}$$

$$\tilde{b}_t+\bar{w}\sum_{s=0}^{T}\beta^s((1+\frac{1}{\eta})(\bar{E}_t\hat{w}_{t+s})+\frac{\bar{\Xi}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{\Xi}_{t+s})-\frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{\tau}_{t+s})$$

$$-\left(\frac{\sigma}{\eta}\bar{w}+(1-\bar{g})\right)\sum_{s=1}^{T}\beta^s\sum_{j=0}^{s-1}\frac{1}{\sigma}(\bar{E}_t i_{t+j}-\bar{E}_t\pi_{t+j+1})$$

$$-\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}\frac{1}{\sigma}\sum_{j=0}^{T-1}(\bar{E}_t i_{t+j}-\bar{E}_t\pi_{t+j+1})-\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}\bar{E}_t i_{t+T}.$$

A.5.2 Optimal pricing decision

Log-linearizing (28) gives

$$\hat{p}_t^*(j)-\hat{p}_t$$
$$=(1-\omega\beta)\tilde{E}_t^j\sum_{s=0}^{T}\omega^s\beta^s\left(\hat{mc}_{t+s}+\hat{p}_{t+s}-\hat{p}_t\right)+(\omega\beta)^{T+1}\tilde{E}_t^j\left(\hat{p}_{t+T}-\hat{p}_t\right),$$

which can be written in terms of inflation expectations as:

$$\hat{p}_t^*(j)-\hat{p}_t=(1-\omega\beta)\left[\hat{mc}_t+\tilde{E}_t^j\sum_{s=1}^{T}\omega^s\beta^s\left(\hat{mc}_{t+s}+\sum_{\tau=1}^{s}\pi_{t+\tau}\right)\right]+(\omega\beta)^{T+1}\tilde{E}_t^j\sum_{\tau=1}^{T}\pi_{t+\tau}. \tag{51}$$

Next, (29) can be log-linearized to yield:

$$\hat{p}_t=\omega\hat{p}_{t-1}+(1-\omega)\int_0^1\hat{p}_t^*(j)dj,$$

from which it follows that:

$$\pi_t=\frac{1-\omega}{\omega}\left(\int_0^1\hat{p}_t^*(j)dj-\hat{p}_t\right). \tag{52}$$

Aggregating (51) and plugging the result in the above expression gives:

$$\pi_t = \frac{(1-\omega)(1-\omega\beta)}{\omega}\left(\hat{mc}_t + \sum_{s=1}^{T}\omega^s\beta^s\hat{E}_t\hat{mc}_{t+s} + \sum_{s=1}^{T}\omega^s\beta^s\sum_{\tau=1}^{s}\hat{E}_t\pi_{t+\tau} + \frac{(\omega\beta)^{T+1}}{1-\omega\beta}\sum_{\tau=1}^{T}\hat{E}_t\pi_{t+\tau}\right). \tag{53}$$

Writing the double sum as a geometric series and combining the outcome with the final term, we can obtain:

$$\pi_t = \frac{(1-\omega)(1-\omega\beta)}{\omega}\sum_{s=0}^{T}\omega^s\beta^s\hat{E}_t\hat{mc}_{t+s} + \frac{(1-\omega)}{\omega}\sum_{s=1}^{T}\omega^s\beta^s\hat{E}_t\pi_{t+s}. \tag{54}$$

A.5.3 Final model

To complete the model, we first log-linearize the market clearing condition (32) and obtain:

$$\hat{Y}_t = (1-\bar{g})\hat{C}_t + \bar{g}\hat{G}_t, \tag{55}$$

and then write wages and marginal costs as:

$$\hat{mc}_t = \hat{w}_t = \eta\hat{H}_t + \sigma\hat{C}_t = (\eta + \frac{\sigma}{1-\bar{g}})\hat{Y}_t - \sigma\frac{\bar{g}}{1-\bar{g}}\hat{G}_t. \tag{56}$$

Finally, we log-linearize the profits of firm j:

$$\hat{\Xi}_t(j) = \frac{1}{1-\bar{mc}}(\hat{p}_t(j) - \hat{p}_t) + \hat{Y}_t(j) - \frac{\bar{mc}}{1-\bar{mc}}\hat{mc}_t, \tag{57}$$

and we aggregate profits over firms:

$$\hat{\Xi}_t = \hat{Y}_t - (\theta-1)\hat{mc}_t, \tag{58}$$

where we use that $\bar{mc} = \frac{\theta-1}{\theta}$.

Using (55) in (50) results in an expression for aggregate output:

$$\begin{aligned}
&\rho\hat{Y}_t = \qquad (59)\\
&\tilde{b}_t + \rho\bar{g}\hat{G}_t + \delta\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{w}_{t+s}) + \frac{\bar{\Xi}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{\Xi}_{t+s}) - \frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{\tau}_{t+s})\\
&-\mu\sum_{s=1}^{T}\beta^s\sum_{j=0}^{s-1}(\bar{E}_t i_{t+j} - \bar{E}_t\pi_{t+j+1})\\
&-\beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}\frac{1}{\sigma}\sum_{j=0}^{T-1}(\bar{E}_t i_{t+j} - \bar{E}_t\pi_{t+j+1}) - \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}\bar{E}_t i_{t+T},
\end{aligned}$$

where

$$\delta = \bar{w}\frac{\eta+1}{\eta}, \tag{60}$$

$$\mu = \frac{\bar{w}}{\eta} + \frac{1-\bar{g}}{\sigma}, \tag{61}$$

and

$$\rho = \frac{1}{1-\bar{g}}\left[\frac{\sigma}{\eta}\bar{w}\frac{1-\beta^{T+1}}{1-\beta} + \frac{1-\bar{g}}{(1-\beta)}\right]. \tag{62}$$

We now assume that agents know, or have learned, about the above relations between aggregate variables (which hold in every period). Therefore, expectations about wages and profits can be substituted for, using (56) and (58). This gives the following system of equations that, together with the specification of monetary and fiscal policy and the government budget constraint, completely describe our model:

$$\begin{aligned} &(\rho-\nu_y)\hat{Y}_t = \qquad (63)\\ &\tilde{b}_t + \rho\bar{g}\hat{G}_t + \nu_g\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{G}_{t+s}) - \frac{\bar{\tau}}{\bar{Y}}\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{\tau}_{t+s}) + \nu_y\sum_{s=1}^{T}\beta^s(\bar{E}_t\hat{Y}_{t+s})\\ &-\mu\sum_{s=1}^{T}\beta^s\sum_{j=0}^{s-1}(\bar{E}_t i_{t+j} - \bar{E}_t\pi_{t+j+1}) - \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}\frac{1}{\sigma}\sum_{j=0}^{T-1}(\bar{E}_t i_{t+j} - \bar{E}_t\pi_{t+j+1}) - \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}\bar{E}_t i_{t+T}, \end{aligned}$$

$$\pi_t = \kappa\sum_{s=0}^{T}\omega^s\beta^s\hat{E}_t\hat{mc}_{t+s} + \frac{(1-\omega)}{\omega}\sum_{s=1}^{T}\omega^s\beta^s\hat{E}_t\pi_{t+s}, \tag{64}$$

with

$$\hat{mc}_t = (\eta + \frac{\sigma}{1-\bar{g}})\hat{Y}_t - \sigma\frac{\bar{g}}{1-\bar{g}}\hat{G}_t, \tag{65}$$

$$\nu_y = \frac{1}{\theta} + \left(\delta - \frac{\theta-1}{\theta}\right)(\eta + \frac{\sigma}{1-\bar{g}}), \tag{66}$$

$$\nu_g = \left(\frac{\theta-1}{\theta} - \delta\right)\frac{\sigma\bar{g}}{1-\bar{g}}, \tag{67}$$

and

$$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}. \tag{68}$$

The government budget constraint (31) is linearized as:

$$\tilde{b}_{t+1} = \frac{\bar{G}}{\bar{Y}\beta}\hat{G}_t - \frac{\bar{\tau}}{\beta\bar{Y}}\hat{\tau}_t + \frac{1}{\beta}\tilde{b}_t. \tag{69}$$

Iterating forward the government budget constraint (69), we have:

$$\bar{E}_t\beta^{T+1}\tilde{b}_{t+T+1} = \bar{E}_t\sum_{i=0}^{T}\beta^i\left(\frac{\bar{G}}{\bar{Y}}\hat{G}_{t+i} - \frac{\bar{\tau}}{\bar{Y}}\hat{\tau}_{t+i}\right) + \tilde{b}_t. \tag{70}$$

Assuming that agents know that all households make decisions according to the same first-order conditions, we can aggregate the first-order conditions (42) and (43) and combine them to substitute for $\tilde{b}_{t+T+1}$ in the above equation. This results in:

$$\tilde{b}_t = \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)}(\hat{C}_t + \sum_{j=0}^{T-1}\frac{1}{\sigma}(\bar{E}_t i_{t+j} - \bar{E}\pi_{t+j+1})) + \beta^{T+1}\frac{1-\bar{g}}{(1-\beta)\sigma}\bar{E}_t i_{t+T} - \bar{E}_t\sum_{i=0}^{T}\beta^i\left(\frac{\bar{G}}{\bar{Y}}\hat{G}_{t+i} - \frac{\bar{\tau}}{\bar{Y}}\hat{\tau}_{t+i}\right). \tag{71}$$

We assume that agents know that the value of current debt depends on the deficit that the government runs in the periods within the horizon, and on how households value debt at the end of the horizon. This implies that Ricardian equivalence theoretically holds in our model. In particular, we can use the above equation to substitute for debt in (63). This simplifies that equation and make future taxes redundant:

$$\begin{aligned}(\tilde{\rho} - \nu_y)\hat{Y}_t &= \tilde{\rho}\bar{g}\hat{G}_t + \tilde{\nu}_g\sum_{s=0}^{T}\beta^s(\bar{E}_t\hat{G}_{t+s}) + \nu_y\sum_{s=1}^{T}\beta^s(\bar{E}_t\hat{Y}_{t+s}) \\ &- \mu\sum_{s=1}^{T}\beta^s\sum_{j=0}^{s-1}(\bar{E}_t i_{t+j} - \bar{E}_t\pi_{t+j+1}),\end{aligned} \tag{72}$$

with

$$\tilde{\rho} = \frac{1}{1-\bar{g}}\left(\frac{\sigma}{\eta}\bar{w} + (1-\bar{g})\right)\frac{1-\beta^{T+1}}{1-\beta}, \tag{73}$$

and:

$$\tilde{\nu}_g = \left(\frac{\theta-1}{\theta} - \delta\right)\frac{\sigma\bar{g}}{1-\bar{g}} - \bar{g}. \tag{74}$$

Finally, log-linearizing (33) gives the Taylor rule:

$$i_t = \phi_1\pi_t. \tag{75}$$

We are left with the derivation of inflation and interest rate expectations in order to obtain a DGP where current output only depends on the path of expected output values and the announced path of government spending within the horizon T.

A.6 Inflation and interest rate expectations

We assume that inflation expectations are consistent with output expectations given by the subjects and with the announced path of government spending. We furthermore assume that agents believe the model to be in steady state after their horizon. Note that this assumption does not influence agents decisions directly, since they only rely on plans and expectations up to the end of the horizon.

We solve for model-consistent inflation expectations by starting with the last period of the horizon. The Philips curve looks like:

$$\hat{E}_t\pi_{t+T} = \kappa\hat{E}_t\hat{mc}_{t+T}. \tag{76}$$

Model-consistent inflation expectations are therefore immediately pinned down by final periods output and government spending expectations. One period earlier, we can write inflation as:

$$\hat{E}_t\pi_{t+T-1} = \kappa\hat{E}_t\hat{mc}_{t+T-1} + \kappa\omega\beta\hat{E}_t\hat{mc}_{t+T} + \frac{1-\omega}{\omega}\omega\beta\hat{E}_t\hat{\pi}_{t+T} = \kappa\hat{E}_t\hat{mc}_{t+T-1} + \kappa\beta\hat{E}_t\hat{mc}_{t+T}. \tag{77}$$

Continuing this process results in the general formula:

$$\hat{E}_t\pi_{t+T-s} = \kappa\sum_{j=0}^{s}\beta^j\hat{E}_t\hat{mc}_{t+T-s+j}. \tag{78}$$

Hence, given an announced path of future government spending, and given an expected path of output given by the subjects, inflation expectations become:

$$\hat{E}_t\pi_{t+T-s} = \kappa\sum_{j=0}^{s}\beta^j(\eta+\frac{\sigma}{1-\bar{g}})\hat{E}_t\hat{Y}_{t+T-s+j} - \sigma\frac{\bar{g}}{1-\bar{g}}\hat{E}_t\hat{G}_{t+T-s+j}, s = 0,1,...,T-1. \tag{79}$$

Using the Taylor rule (75), nominal interest rate expectations can then be written as:

$$\hat{E}_t i_{t+T-s} = \phi_1\kappa\sum_{j=0}^{s}\beta^j(\eta+\frac{\sigma}{1-\bar{g}})\hat{E}_t\hat{Y}_{t+T-s+j} - \phi_1\sigma\frac{\bar{g}}{1-\bar{g}}\hat{E}_t\hat{G}_{t+T-s+j}, s = 0,1,...,T-1. \tag{80}$$

Plugging these two equations (79) and (80) into (72) shows that the evolution of aggregate output only depends on the expected paths of output and government spending within the planning horizon, and that expectations of all other variables are no longer explicit, hence providing a DGP for our experimental economies.

A.7 Parameterization

β	Discount factor	0.99
σ	Inverse elasticity of intertemporal substitution	2
θ	substituability across goods	6
ω	Calvo probability	0.85
η	labor elasticity	2
ϕ_1	Taylor rule coefficient	1.5
$\bar{g}$	steady state government spending to output ratio	0.25

Table 7: Parameter values of the model

B Experimental results

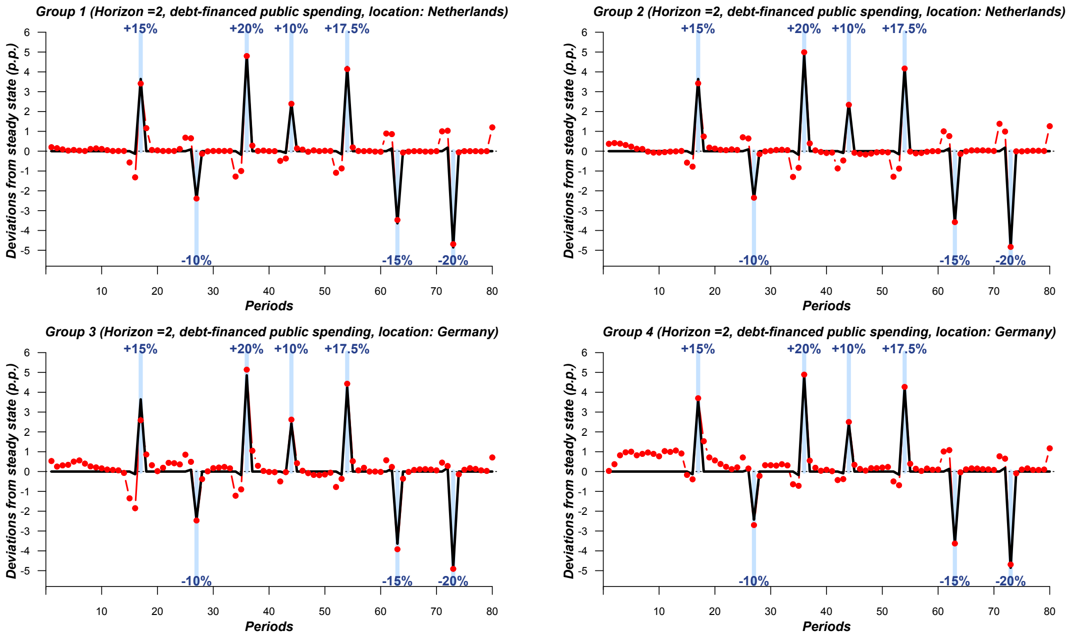


Figure 6: The four experimental groups with **T** = **2** and **debt-financed** spending

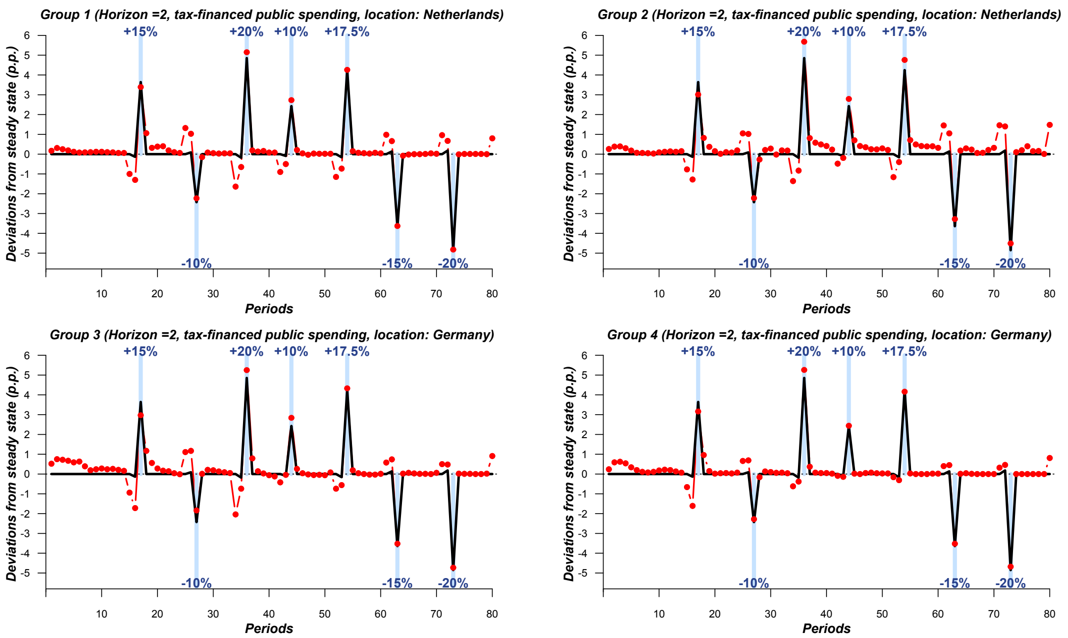


Figure 7: The four experimental groups with **T** = **2** and **tax-financed** spending

Notes: Each graphic represents the output in the experimental economies (red line with dots), the output under perfect foresight (solid black line), the frequency of each fiscal shock (gray-shaded area) and their size.

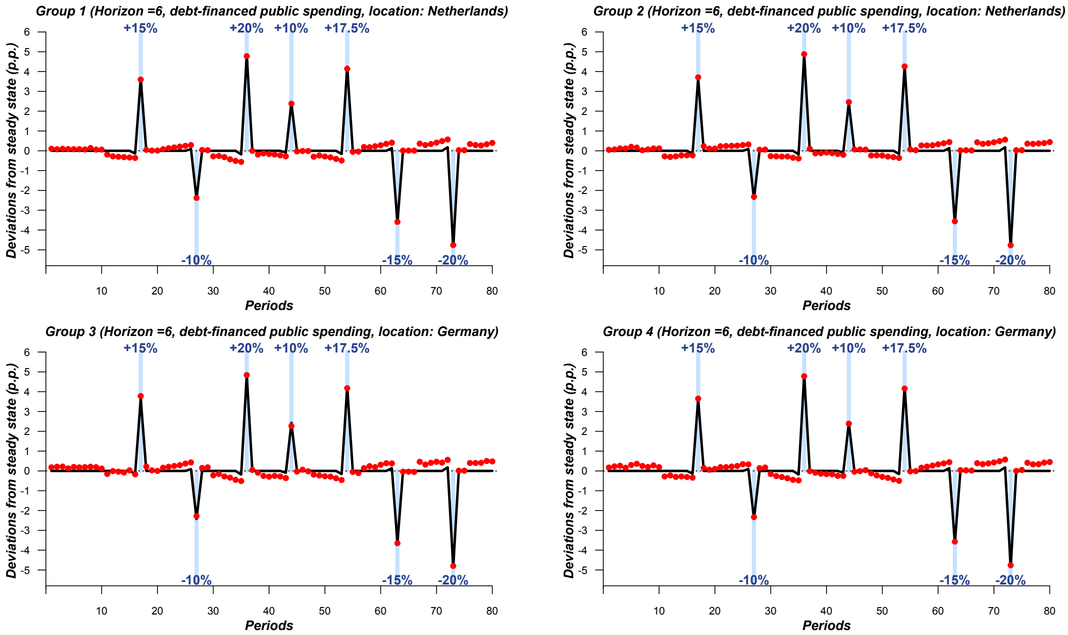


Figure 8: The four experimental groups with $\mathbf{T = 6}$ and **debt-financed** spending

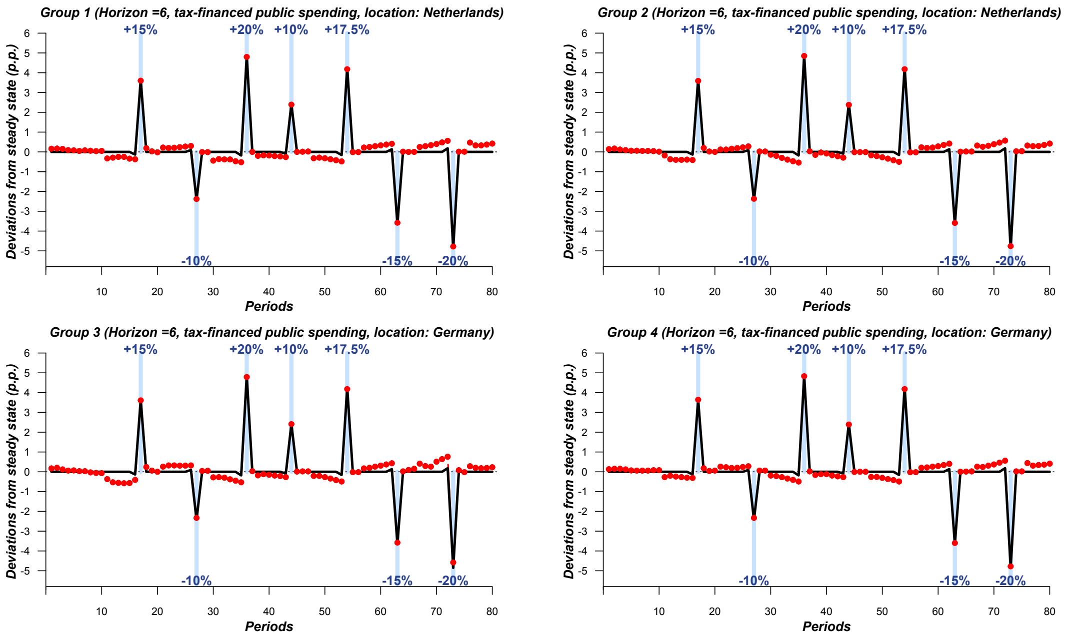


Figure 9: The four experimental groups with $\mathbf{T = 6}$ and **tax-financed** spending

Notes: Each graphic represents the output in the experimental economies (red line with dots), the output under perfect foresight (solid black line), the frequency of each fiscal shock (gray-shaded area) and their size.

Experimental economies with $T = 6$ only

Panel A. Dependent variable: **aggregate *forecast errors***

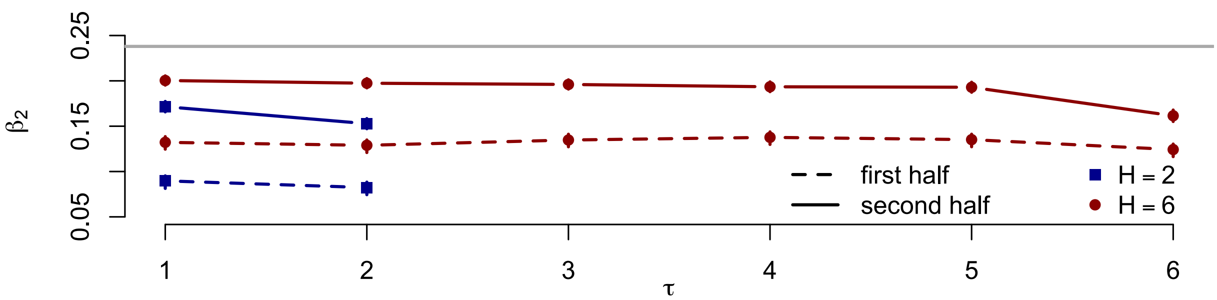
		$\tau = 2$			$\tau = 4$			$\tau = 5$	
	OLS	FE	FE	OLS	FE	FE	OLS	FE	FE
$\hat{b}$	0.119***	0.119***	0.119***	0.055***	0.056***	0.055***	0.059***	0.059***	0.059***
(sd)	(0.036)	(0.018)	(0.019)	(0.010)	(0.015)	(0.010)	(0.013)	(0.015)	(0.015)
y_{t-2}			-0.014			-0.020			-0.011
(sd)			(0.019)			(0.020)			(0.017)
# economies with sign. $b > 0$	6	6	6	6	6	6	7	7	7
# economies	8	8	8	8	8	8	8	8	8

Panel B. Dependent variable: **individual *forecast errors***

		$\tau = 2$			$\tau = 4$			$\tau = 5$	
	OLS	FE	FE	OLS	FE	FE	OLS	FE	FE
$\hat{b}$	-0.132***	-0.132***	-0.133***	-0.097***	-0.098***	-0.099***	-0.122***	-0.122***	-0.122***
(sd)	(0.020)	(0.006)	(0.006)	(0.030)	(0.009)	(0.009)	(0.035)	(0.009)	(0.009)
y_{t-2}			-0.017**			-0.030**			-0.006
(sd)			(0.008)			(0.011)			(0.011)

Notes: see Table 4.

Table 8: Complementary results: two-, four-and five-step-ahead forecasts



Notes:

Figure 10: Forecast reactions to news at various horizons over the two halves of the experiment

C Lab material

In what follows, in the following order:

- the instructions for $T = 2$ in English for the sessions run in Amsterdam;
- the instructions for $T = 6$ in English for the sessions run in Amsterdam;
- the instruction, including the GUI, for $T = 2$ translated in German for the sessions run in Bamberg;
- the instructions including the GUI, for $T = 6$ translated in German for the sessions run in Bamberg;
- the post-experiment questionnaire in English (Section C.1);
- the GUI in English for $T = 2$ (Fig. 11);
- the GUI in English for $T = 6$ (Fig. 12);

Instructions

Welcome to this experiment! The experiment is anonymous, the data from your choices will only be linked to your station ID, never to your name. You will be paid privately once all participants have finished the experiment. We reserve the right to improve your payment in your favor if average payoffs in your group are lower than expected. On your desk there is a calculator, which you can use during the experiment.

During the experiment you are not allowed to communicate with other participants. If you have any question at any time, please raise your hand and someone will come to your desk.

Information about the experimental economy

The experiment is based on a simulation that approximates fluctuations in the real economy. The economy you are participating in is mainly described by **four variables: total output, government expenditures, taxes and government debt**. Your task is to serve as a **professional forecaster** and provide real-time **forecasts about future output** in this simulated economy. There are **nine other forecasters** like you in the economy. The group composition will not change during the experiment. All participants have the same task.

The instructions will now explain what output, government expenditures, taxes and government debt are and how they move around in this economy, as well as how they depend on the forecasts of all forecasters in the economy.

The values of output, government expenditures, taxes and government debt will be given in ***percentage points***, a measurement often used in descriptions of economies. All values can be **positive**, **negative**, or **zero** at any point in time, they simply indicate whether an economic variable is **higher**, **lower** or exactly **at its *normal*, or usual, level**. For instance, a value of 1 % indicates that the variable is 1 % above its normal level, a value of -8.5 % indicates that the variable is 8.5 % below its normal value.

Output is the total production of goods in the economy. All production is either consumed by households (three quarters of total output) or by the government (one quarter of total output). Hence, output is the sum of the households' consumption and government expenditures.

Output is displayed in the top graph on your screen (see the separate sheet on your desk). On the horizontal axis are the time periods; the vertical axis is in percentages and shows how output (pink line) deviates from its normal level. The usual level is **0 %**. However, output can be **higher** (that is, **above 0 %**), or **lower** (that is, **below 0 %**) for two reasons.

The first reason is because households' **consumption can change as a result of your forecasts and the forecasts of the other participants** in the economy. Specifically, the households in the economy use the **average output forecasts** across all bureaus like yours to decide about their consumption.

The second reason why output may deviate from its normal level is because government expenditures may change. The government finances its expenditures either by immediately **raising taxes** on households or **by issuing government debt**. The government has usually no debt, but if it accumulates some, debt **always has to be repaid at some point in the future through taxes on households**: the government cannot raise its debt without limits. Hence, **how government expenditures are exactly financed does not matter directly for output**. However, whether government expenditures are financed through debt or taxes may affect output **if it affects your forecasts or the forecasts of the other participants** in the economy.

The second plot displays the evolution of government debt (black line), taxes (blue lines) and government expenditures (brown line), see again the separate sheet on your desk. Again, the y-axis indicates percentage points, the normal level of any of those variables is 0 %, but they may be higher or lower.

If government expenditures increase, **output mechanically increases**. **However**, as output is **only partly** composed of government expenditures, an increase in government expenditures is **not** likely to affect output **with the same magnitude**.

Moreover, an increase in government expenditures also **decreases consumption**, because households either have to pay more taxes now (if the expenditures are financed by immediately raising taxes) or have to save to **pay more taxes** in the future (if the expenditures are financed by debt).[1] This negative effect on consumption is smaller than the mechanical increase in output following an increase in government expenditures.

Therefore, ***output is likely to increase when government expenditures increase***.

[1] Additionally an increase in the expenditures of the government will lead to an increase in the real interest rate which also induces households to save more and consume less.

However, the **overall effect of an increase in government expenditures on output also depends on your output forecasts and the forecasts of the other participants** in the economy. Precisely, as explained above, ***an increase in the average forecast increases output***.

Those different effects of an increase in government expenditures on output are summarized in Figure 1 below.

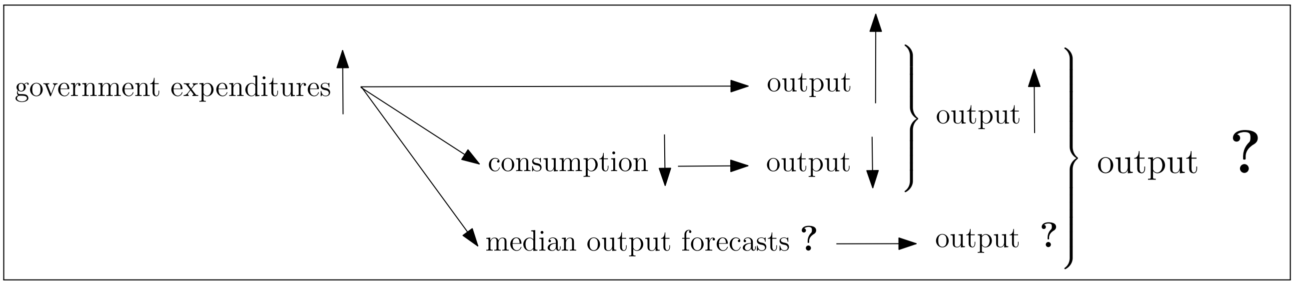


Figure 1: Different effects of an increase in government expenditures on output.

The same reasoning applies conversely in case of a **decrease in government expenditures**: **output is likely to decrease but the overall effect depends on the average output forecast** (***a decrease in the average forecast decreases output***).

Your task

The experiment will last for 80 periods. In each period, you have to submit **two output forecasts: one for each of the next two periods**. You have to enter your forecasts using the **two sliders** under the plots (see again the separate sheet). Once you are satisfied with your forecasts, press the '`Submit`' button.

In period 1, you have to enter output forecasts for periods 2 and 3. Once every participant has submitted his/her forecasts, output in period 1 is displayed and the experiment moves to period 2. You then have to submit output forecasts for periods 3 and 4. In period 3, you have to forecast output for periods 4 and 5, etc.

This implies that **you have to submit output forecasts two times for the same period**. For instance, you have to forecast output for period 5 in period 3 and then again in period 4. In the top graph, next to output, we also report your output forecasts. The graph does not report all your previous forecasts, but only your rewarded forecast for previous periods (black dots, see below how they are selected), and your most recent forecast for the current period and future periods. This information runs until the previous period.

The government will **always announce government expenditures, taxes and debt for the next two periods and will stick to its announcements**. This will be displayed in the bottom graph. This way, you always have information about those variables for the periods for which you have to forecast output.

On the screen, in period 1, the two sliders are initialized at the normal output, namely 0 %. In all other periods, the first slider is initialized at the value that you have forecast in the previous period. The second slider, which corresponds to the new period for which you have not made any forecast yet, will be initialized at the normal output 0 %. You can try out different forecast values by adjusting the sliders, and the corresponding forecast output path over the next two periods (red line) will be automatically adjusted on the first plot. If you wish to go back to the values at which the sliders were initialized, you may click the '`Reset`' button. Note that you can zoom in the plots by clicking and dragging (to select the area you want to zoom in).

In the upper-right corner of your screen, a timer indicates the remaining time that you have to submit your forecasts. When the time is up, the message "The time is up!!" appears. You will still be able to submit your forecasts when the time is up but in your own interest, we urge you to quickly submit your forecasts **so as not to delay the experiment**. At the beginning of the experiment, you will be given a bit more time to get familiar with your task.

To sum up, in any period, you will have the following information:

- The history of output and your past rewarded output forecasts (up to the previous period);
- The history and future paths of government expenditures, taxes and debt (up to two periods in the future);

All this information may be relevant to form your forecasts, but it is up to you to determine how to use it. You can also see this information in a table by clicking on the '`Table`' button at the top right of the screen. In later periods, you may have to scroll down the table to see all the previous periods. In the table, you also see the number of points you have already earned. We now explain how these are computed.

Information on payment

During the experiment you collect points. **The amount of points depends on the accuracy of your forecasts**. The accuracy is measured by the absolute distance between the realized value of output in a particular period and one of your forecasts for that period. This rewarded forecast is randomly chosen out of the two forecasts that you submitted for that period (the two are equally likely to be chosen). The rewarded forecasts are the ones that are plotted on the upper graph (black dots) and reported in the table (in the column 'Rewarded forecast').

Figure 2 below presents the amount of points that you make as a function of your absolute forecast error, which is computed as $\frac{100}{1+\text{absolute error}}$. The maximum amount of points per period is 100 and is earned in case of a perfect prediction (zero error). The table on your screen shows the points that you have earned in each period. At the end of the experiment, your total amount of points is transformed into euros and paid to you privately. **One euro** corresponds to **300 points**.

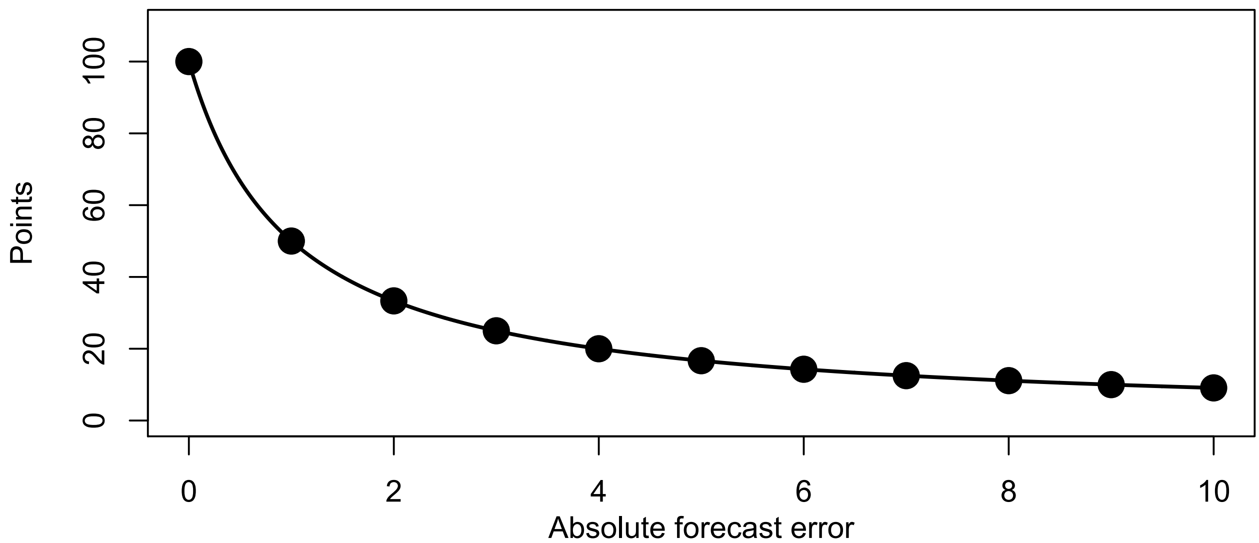


Figure 2: Relation between your forecast errors and the points you earn.

Example: You enter period 11, output in period 10 is just revealed and is 1.5 %. You will be rewarded for a forecast that you submitted for output in period 10. You submitted two forecasts for period 10: one in period 8 and one in period 9. Each forecast has a $\frac{1}{2}$ probability of being chosen for reward. Imagine that the forecast submitted in period 8 is chosen (and then displayed on the top graph on your screen) and you had submitted -4 %. Since output turned out to be 1.5 %, your error is $\mid 1.5 - (-4) \mid = 5.5$ percentage points, and the figure indicates that you earn $\frac{100}{1+5.5} = 15$ points.

Instructions

Welcome to this experiment! The experiment is anonymous, the data from your choices will only be linked to your station ID, never to your name. You will be paid privately once all participants have finished the experiment. We reserve the right to improve your payment in your favor if average payoffs in your group are lower than expected. On your desk there is a calculator, which you can use during the experiment.

During the experiment you are not allowed to communicate with other participants. If you have any question at any time, please raise your hand and someone will come to your desk.

Information about the experimental economy

The experiment is based on a simulation that approximates fluctuations in the real economy. The economy you are participating in is mainly described by **four variables: total output, government expenditures, taxes and government debt**. Your task is to serve as a **professional forecaster** and provide real-time **forecasts about future output** in this simulated economy. There are **nine other forecasters** like you in the economy. The group composition will not change during the experiment. All participants have the same task.

The instructions will now explain what output, government expenditures, taxes and government debt are and how they move around in this economy, as well as how they depend on the forecasts of all forecasters in the economy.

The values of output, government expenditures, taxes and government debt will be given in ***percentage points***, a measurement often used in descriptions of economies. All values can be **positive**, **negative**, or **zero** at any point in time, they simply indicate whether an economic variable is **higher**, **lower** or exactly **at its *normal*, or usual, level**. For instance, a value of 1 % indicates that the variable is 1 % above its normal level, a value of -8.5 % indicates that the variable is 8.5 % below its normal value.

Output is the total production of goods in the economy. All production is either consumed by households (three quarters of total output) or by the government (one quarter of total output). Hence, output is the sum of the households' consumption and government expenditures.

Output is displayed in the top graph on your screen (see the separate sheet on your desk). On the horizontal axis are the time periods; the vertical axis is in percentages and shows how output (pink line) deviates from its normal level. The usual level is **0 %**. However, output can be **higher** (that is, **above 0 %**), or **lower** (that is, **below 0 %**) for two reasons.

The first reason is because households' **consumption can change as a result of your forecasts and the forecasts of the other participants** in the economy. Specifically, the households in the economy use the **average output forecasts** across all bureaus like yours to decide about their consumption.

The second reason why output may deviate from its normal level is because government expenditures may change. The government finances its expenditures either by immediately **raising taxes** on households or **by issuing government debt**. The government has usually no debt, but if it accumulates some, debt **always has to be repaid at some point in the future through taxes on households**: the government cannot raise its debt without limits. Hence, **how government expenditures are exactly financed does not matter directly for output**. However, whether government expenditures are financed through debt or taxes may affect output **if it affects your forecasts or the forecasts of the other participants** in the economy.

The second plot displays the evolution of government debt (black line), taxes (blue lines) and government expenditures (brown line), see again the separate sheet on your desk. Again, the y-axis indicates percentage points, the normal level of any of those variables is 0 %, but they may be higher or lower.

If government expenditures increase, **output mechanically increases**. **However**, as output is **only partly** composed of government expenditures, an increase in government expenditures is **not** likely to affect output **with the same magnitude**.

Moreover, an increase in government expenditures also **decreases consumption**, because households either have to pay more taxes now (if the expenditures are financed by immediately raising taxes) or have to save to **pay more taxes** in the future (if the expenditures are financed by debt).[1] This negative effect on consumption is smaller than the mechanical increase in output following an increase in government expenditures.

Therefore, ***output is likely to increase when government expenditures increase***.

[1] Additionally an increase in the expenditures of the government will lead to an increase in the real interest rate which also induces households to save more and consume less.

However, the **overall effect of an increase in government expenditures on output also depends on your output forecasts and the forecasts of the other participants** in the economy. Precisely, as explained above, ***an increase in the average forecast increases output***.

Those different effects of an increase in government expenditures on output are summarized in Figure 1 below.

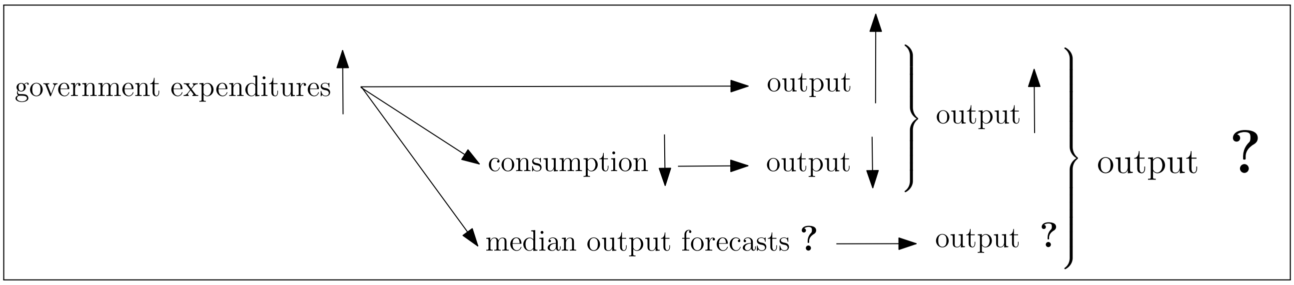


Figure 1: Different effects of an increase in government expenditures on output.

The same reasoning applies conversely in case of a **decrease in government expenditures**: **output is likely to decrease but the overall effect depends on the average output forecast** (***a decrease in the average forecast decreases output***).

Your task

The experiment will last for 80 periods. In each period, you have to submit **six output forecasts: one for each of the next six periods**. You have to enter your forecasts using the **six sliders** under the plots (see again the separate sheet). Once you are satisfied with your forecasts, press the '`Submit`' button.

In period 1, you have to enter output forecasts for periods 2, 3, 4, 5, 6 and 7. Once every participant has submitted his/her forecasts, output in period 1 is displayed and the experiment moves to period 2. You then have to submit output forecasts for periods 3, 4, 5, 6, 7 and 8. In period 3, you have to forecast output for periods 4 to 9, etc.

This implies that **you have to submit output forecasts six times for the same period**. For instance, you have to forecast output in period 15 for the first time in period 9, then again in period 10, 11, 12, 13 and finally in period 14. In the top graph, next to output, we also report your output forecasts. The graph does not report all your previous forecasts, but only your rewarded forecast for previous periods (black dots, see below how they are selected), and your most recent forecast for the current period and future periods. This information runs until the previous period.

The government will **always announce government expenditures, taxes and debt for the next six periods and will stick to its announcements**. This will be displayed in the bottom graph. This way, you always have information about those variables for the periods for which you have to forecast output.

On the screen, in period 1, the six sliders are initialized at the normal output, namely 0 %. In all other periods, the first five sliders are initialized at the values that you have forecast in the previous period. The sixth slider, which corresponds to the new period for which you have not made any forecast yet, will be initialized at the normal output 0 %. You can try out different forecast values by adjusting the sliders, and the corresponding forecast output path over the next six periods (red line) will be automatically adjusted on the first plot. If you wish to go back to the values at which the sliders were initialized, you may click the '`Reset`' button. Note that you can zoom in the plots by clicking and dragging (to select the area you want to zoom in).

In the upper-right corner of your screen, a timer indicates the remaining time that you have to submit your forecasts. When the time is up, the message "The time is up!!" appears. You will still be able to submit your forecasts when the time is up but in your own interest, we urge you to quickly submit your forecasts **so as not to delay the experiment**. At the beginning of the experiment, you will be given a bit more time to get familiar with your task.

To sum up, in any period, you will have the following information:

- The history of output and your past rewarded output forecasts (up to the previous period);
- The history and future paths of government expenditures, taxes and debt (up to six periods in the future);

All this information may be relevant to form your forecasts, but it is up to you to determine how to use it. You can also see this information in a table by clicking on the '`Table`' button at the top right of the screen. In later periods, you may have to scroll down the table to see all the previous periods. In the table, you also see the number of points you have already earned. We now explain how these are computed.

Information on payment

During the experiment you collect points. **The amount of points depends on the accuracy of your forecasts**. The accuracy is measured by the absolute distance between the realized value of output in a particular period and one of your forecasts for that period. This rewarded forecast is randomly chosen out of the six forecasts that you submitted for that period (the six are equally likely to be chosen). The rewarded forecasts are the ones that are plotted on the upper graph (black dots) and reported in the table (in the column 'Rewarded forecast').

Figure 2 below presents the amount of points that you make as a function of your absolute forecast error, which is computed as $\frac{100}{1+\text{absolute error}}$. The maximum amount of points per period is 100 and is earned in case of a perfect prediction (zero error). The table on your screen shows the points that you have earned in each period. At the end of the experiment, your total amount of points is transformed into euros and paid to you privately. **One euro** corresponds to **300 points**.

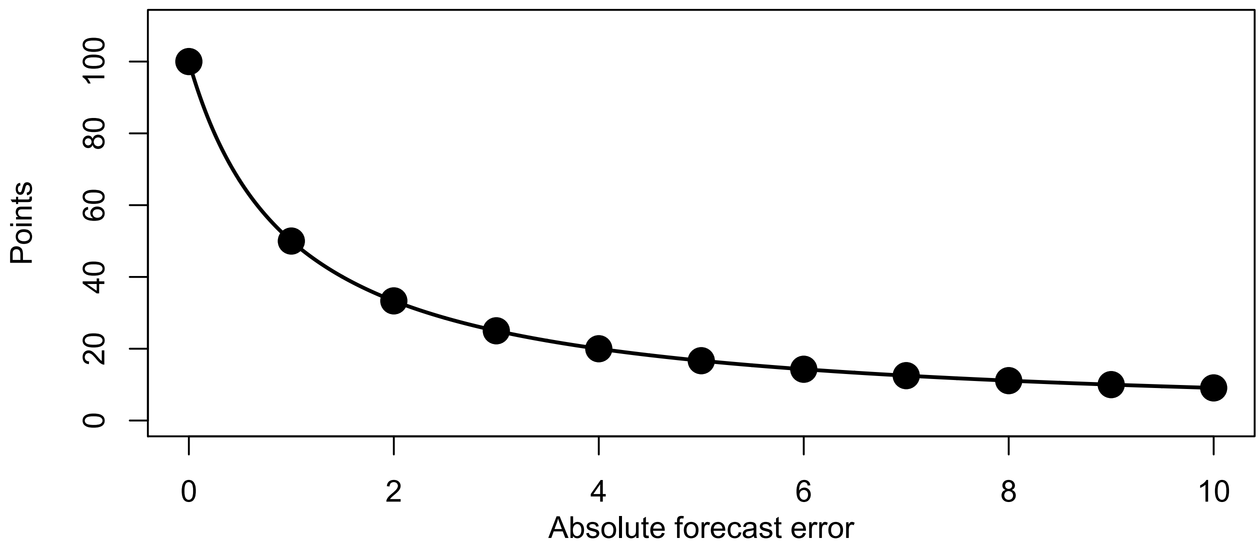


Figure 2: Relation between your forecast errors and the points you earn.

Example: You enter period 11, output in period 10 is just revealed and is 1.5 %. You will be rewarded for a forecast that you submitted for output in period 10. You submitted six forecasts for period 10: one submitted in period 4, one in period 5, one in period 6, one in period 7, one in period 8 and one in period 9. Each forecast has a $\frac{1}{6}$ probability of being chosen for reward. Imagine that the forecast submitted in period 8 is chosen (and then displayed on the top graph on your screen) and you had submitted -4 %. Since output turned out to be 1.5 %, your error is $\mid 1.5 - (-4) \mid = 5.5$ percentage points, and the figure indicates that you earn $\frac{100}{1+5.5} = 15$ points.

Anleitung

Herzlich willkommen zu diesem Experiment! Dieses Experiment ist anonym und die Daten aus Ihren Entscheidungen werden nur mit der Stations-ID verknüpft und nicht mit ihrem Namen. Sie werden am Ende, nachdem alle Teilnehmer das Experiment abgeschlossen haben, bezahlt. Wir behalten uns das Recht vor, Ihre Zahlung zu Ihren Gunsten zu verbessern, wenn die durchschnittlichen Auszahlungen in Ihrer Gruppe niedriger sind als erwartet. Auf Ihrem Tisch liegt ein Taschenrechner, den Sie während des Eperiments benutzen können.

Während des Experiments dürfen Sie nicht mit anderen Teilnehmern kommunizieren. Sobald Sie zu irgendeinem Zeitpunkt eine Frage haben sollten, heben Sie bitte die Hand und es wird jemand zu Ihrem Tisch kommen.

Informationen über die fiktive Ökonomie

Das Experiment basiert auf einer Simulation, welche die Schwankungen der realen Ökonomie approximiert. Die Ökonomie, an der Sie teilnehmen, wird hauptsächlich durch **vier Variablen** beschrieben: **Gesamtproduktion, Staatsausgaben, Steuern und Staatsschulden**. Ihre Aufgabe ist es als **professioneller Analyst** zu agieren und Echtzeit-Prognosen über die **zukünftige Produktion** in dieser simulierten Ökonomie zu erstellen. Es gibt **neun weitere Analysten** wie Sie in dieser Ökonomie. Die Gruppenzusammensetzung wird sich im Laufe des Experiments nicht verändern. Alle Teilnehmer haben dieselbe Aufgabe.

Die Anleitung wird nun erklären, was Produktion, Staatsausgaben, Steuern und Staatsverschuldung sind, wie sie sich in dieser Ökonomie entwickeln und wie sie von den Prognosen aller Analysten in der Ökonomie abhängen.

Die Werte von Produktion, Staatsausgaben, Steuern und Staatsschulden werden in ***Prozentpunkten*** angegeben, wie es häufig in den Beschreibungen von Volkswirtschaften der Fall ist. Alle Werte können zu jedem Zeitpunkt **positiv**, **negativ** oder **Null** sein und geben lediglich an, ob eine wirtschaftliche Variable **höher**, **niedriger** oder genau **auf ihrem *normalen* bzw. üblichen Niveau ist**. Ein Wert von 1 % bedeutet beispielsweise, dass die Variable 1 % über ihrem normalen Niveau liegt, ein Wert von -8,5 % wiederum bedeutet, dass die Variable 8,5 % unter ihrem normalen Niveau liegt.

Die Produktion ist die Gesamtproduktion von Gütern in der Ökonomie. Die gesamte Produktion wird von den Haushalten (drei Viertel der Gesamtproduktion) und vom Staat (ein Viertel der Gesamtproduktion) konsumiert. Die Produktion ist somit die Summe aus dem Konsum der Haushalte und den Staatsausgaben.

Die Produktion wird im oberen Diagramm auf Ihrem Bildschirm angezeigt (siehe separates Blatt auf Ihrem Schreibtisch). Auf der horizontalen Achse befinden sich die Zeitperioden; die vertikale Achse ist in Prozentpunkten angegeben und zeigt, wie die Produktion (magenta Linie) von ihrem normalen Niveau abweicht. Das normale Niveau beträgt **0 %**. Die Produktion kann jedoch aus zwei Gründen **höher** (d.h. **über 0 %**) oder **niedriger** (d.h. **unter 0 %**) sein.

Der erste Grund ist, dass sich der **Konsum der Haushalte aufgrund Ihrer Prognosen und der Prognosen der anderen Wirtschaftsteilnehmer (Analysten) ändern kann**. Konkret nutzen die Haushalte in der Ökonomie die **durchschnittlichen Produktionsprognosen** der Analysten (auch Ihre Prognosen), um über ihren Konsum zu entscheiden.

Der zweite Grund, warum die Produktion von ihrem normalen Niveau abweichen kann, liegt in der Tatsache, dass sich die Staatsausgaben ändern können. Der Staat finanziert seine Ausgaben entweder durch die sofortige **Erhöhung der Steuern** auf die Haushalte oder durch **staatliche Schuldverschreibungen**. Der Staat hat in der Regel keine Schulden, aber wenn er welche anhäuft, muss er die Schulden **immer zu irgendeinem Zeitpunkt in der Zukunft durch die Erhebung von Steuern auf die Haushalte finanzieren**: Der Staat kann seine Schulden nicht unbegrenzt erhöhen. Ob die Staatsausgaben schulden- oder steuerfinanziert sind, spielt für die Produktion keine Rolle, es sei denn es **beeinflusst Ihre Prognosen oder die Prognosen der anderen Analysten**.

Die zweite Grafik zeigt die Entwicklung der Staatsverschuldung (schwarze Linie), der Steuern (blaue Linie) und der Staatsausgaben (braune Linie). Vgl. das separate Blatt auf Ihrem Tisch. Auch hier ist die y-Achse in Prozentpunkten angegeben und das normale Niveau jeder dieser Variablen liegt bei 0 %. Die Variablen können aber höher oder niedriger als 0 % werden.

Wenn die Staatsausgaben steigen, **wird die Produktion ebenfalls automatisch ansteigen**. Da sich die Produktion jedoch **nur zum Teil** aus den Staatsausgaben zusammensetzt, würde ein Anstieg der Staatsausgaben die Produktion **nicht in gleichem Maße** beeinflussen.

Darüber hinaus **verringert** ein Anstieg der Staatsausgaben auch **den Konsum**, da die Haushalte entweder sofort mehr Steuern zahlen müssen (wenn die Ausgaben durch sofortige

Steuererhöhungen finanziert werden) oder in Zukunft **mehr Steuern zahlen** müssen (wenn die Ausgaben durch Schulden finanziert werden).[1] Dieser negative Effekt auf den Konsum ist kleiner als der automatische Anstieg der Produktion, der infolge einer Staatsausgabenerhöhung eintritt.

Daher ***dürfte die Produktion steigen, wenn die Staatsausgaben steigen***. Der **Gesamteffekt eines Anstiegs der Staatsausgaben auf die Produktion hängt jedoch auch von Ihren Produktionsprognosen und den Prognosen der anderen Analysten** ab. Wie bereits oben erläutert, ***führt eine Erhöhung der durchschnittlichen Prognose zu einer Steigerung der Produktion***.

Diese verschiedenen Effekte eines Anstiegs der Staatsausgaben auf die Produktion sind in Abbildung 1 unten zusammengefasst.

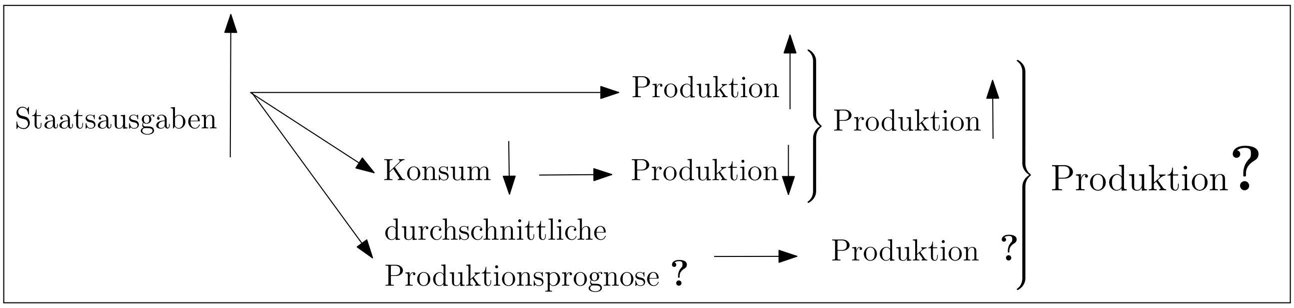


Figure 1: Unterschiedliche Effekte eines Anstiegs der Staatsausgaben auf die Produktion.

Die gleiche Argumentation gilt umgekehrt im Falle eines **Rückgangs der Staatsausgaben**: **Die Produktion dürfte sinken, aber der Gesamteffekt hängt von der durchschnittlichen Produktionsprognose ab** (***ein Rückgang der durchschnittlichen Prognose verringert die Produktion***).

Ihre Aufgabe

Das Experiment dauert 80 Perioden. In jeder Periode müssen Sie **zwei Produktionsprognosen abgeben: eine für jede der nächsten zwei Perioden**. Sie müssen Ihre Prognosen mit den **zwei Reglern**, welche sich unter den Diagrammen befinden, erfassen (siehe ebenso separates Blatt). Wenn Sie mit Ihren Prognosen zufrieden sind, klicken Sie auf die Schaltfläche „`Absenden`".

[1] Zusätzlich wird eine Erhöhung der Ausgaben der Regierung zu einer Erhöhung des realen Zinssatzes führen, was wiederum die Haushalte dazu veranlasst, mehr zu sparen und weniger zu konsumieren.

In Periode 1 müssen Sie Produktionsprognosen für die Perioden 2 und 3 abgeben. Sobald jeder Teilnehmer seine/ihre Prognosen abgegeben hat, wird die Produktion in Periode 1 angezeigt und das Experiment geht in Periode 2 über. Anschließend müssen Sie die Produktionsprognosen für die Perioden 3 und 4 abgeben. In Periode 3 müssen Sie die Produktion für die Perioden 4 und 5 prognostizieren, und so weiter.

Dies bedeutet, dass **Sie insgesamt zwei Produktionsprognosen für den gleichen Zeitraum abgeben müssen**. Beispielsweise müssen Sie die Produktion in Periode 5 zum ersten Mal in Periode 3 und dann wieder in Periode 4 prognostizieren. Im oberen Diagramm zeigen wir neben der Produktion auch Ihre Produktionsprognosen an. Jedoch zeigt die Grafik nicht alle Ihre bisher abgegebenen Prognosen an. Für frühere Perioden werden die jeweils für die Auszahlung herangezogene Prognose angezeigt (schwarze Punkte; sehe unten, wie diese Prognosen ausgewählt werden). Für die aktuelle Periode und die zukünftigen Perioden wird Ihre jüngste Prognose angezeigt.

Die Regierung wird **immer ihre Staatsausgaben, Steuern und Staatsschulden für die nächste zwei Perioden bekannt geben und diese Ankündigungen umsetzen**. Diese Ankündigungen werden im unteren Diagramm angezeigt. Dadurch haben Sie immer Informationen über die Variablen der entsprechenden Perioden, für die Sie die Produktion prognostizieren müssen.

Auf dem Bildschirm werden in Periode 1 die zwei Regler auf das normale Niveau, nämlich 0 %, initialisiert. In allen anderen Perioden wird der erste Regler mit dem Wert initialisiert, den Sie in der Vorperiode prognostiziert haben. Der zweite Regler, welcher der neuen Periode entspricht, für die Sie noch keine Prognose abgegeben haben, wird mit dem Ausgangswert von 0 % initialisiert. Sie können verschiedene Prognosewerte durch Verschieben des Reglers ausprobieren, wobei die entsprechende Prognose über die Produktionsentwicklung der nächsten sechs Perioden (rote Linie) in der ersten Grafik automatisch angepasst wird. Wenn Sie die Werte wieder auf ihre Initialisierungswerte zurücksetzen möchten, können Sie auf die Schaltfläche „`Reset`“ klicken. Beachten Sie, dass Sie die Diagramme durch Klicken und Ziehen vergrößern können (um den gewünschten Ausschnitt auszuwählen und zu vergrößern).

In der oberen rechten Ecke Ihres Bildschirms wird Ihnen die verbleibende Zeit angezeigt, die Sie für die Eingabe Ihrer Prognosen noch zur Verfügung haben. Wenn die Zeit abgelaufen ist, erscheint die Meldung „Die Zeit ist vorbei!!!“. Sie können Ihre Prognosen auch dann noch abgeben, wenn die Zeit abgelaufen ist, aber in Ihrem eigenen Interesse bitten wir Sie, Ihre Prognosen schnell einzureichen, **um das Experiment nicht zu verzögern**. Zu Beginn des Experiments haben Sie etwas mehr Zeit, um sich mit Ihrer Aufgabe vertraut zu machen.

Zusammenfassend haben Sie in jeder Periode die folgenden Informationen:

- Die Entwicklung der Produktion und Ihre in der Vergangenheit für die Auszahlung herangezogenen Produktionsprognosen (bis zur letzten Periode);
- Die bisherige und zukünftige Entwicklung der Staatsausgaben, Steuern und Staatsschulden (bis zu zwei Perioden in der Zukunft);

All diese Informationen können für die Erstellung Ihrer Prognosen relevant sein, aber es liegt an Ihnen, wie Sie diese verwenden. Sie können diese Informationen auch in einer Tabelle sehen, indem Sie auf die Schaltfläche „`Table`“ oben rechts auf dem Bildschirm klicken. Im späteren Verlauf müssen Sie möglicherweise in der Tabelle nach unten blättern, um alle früheren Perioden anzuzeigen. In der Tabelle sehen Sie auch die Anzahl der Punkte, die Sie bereits gesammelt haben. Wir erklären nun, wie diese berechnet werden.

Informationen zur Auszahlung

Während des Experiments sammeln Sie Punkte. **Die Anzahl der Punkte hängt von der Genauigkeit Ihrer Prognosen ab**. Die Genauigkeit wird durch den absoluten Abstand zwischen dem tatsächlichen Wert der Produktion und einem Ihrer Prognosewerte für diesen Zeitraum gemessen. Diese für die Auszahlung herangezogene Produktionsprognose wird nach dem Zufallsprinzip aus den zwei Prognosen, die Sie für diesen Zeitraum abgegeben haben, ausgewählt (beide Prognosewerte haben die gleiche Wahrscheinlichkeit ausgewählt zu werden). Die für die Auszahlung herangezogenen Prognosen („belohnten Prognosen“) sind diejenigen, die in der oberen Grafik (schwarze Punkte) dargestellt und in der Tabelle ausgewiesen werden.

Abbildung 2 zeigt die Anzahl der Punkte, die Sie in Abhängigkeit von Ihrem absoluten Prognosefehler erhalten, welcher durch $\frac{100}{1+\text{absoluter Fehler}}$ berechnet wird. Die maximale Anzahl von Punkten pro Periode beträgt 100, welche Sie bei perfekter Prognose erhalten (kein Prognosefehler). Die Tabelle auf Ihrem Bildschirm zeigt die Punkte, die Sie in jeder Periode erhalten haben. Am Ende des Experiments wird Ihr Gesamtpunktestand in Euro umgewandelt und Ihnen anschließend ausgezahlt. **Ein Euro** entspricht hierbei **300 Punkten**.

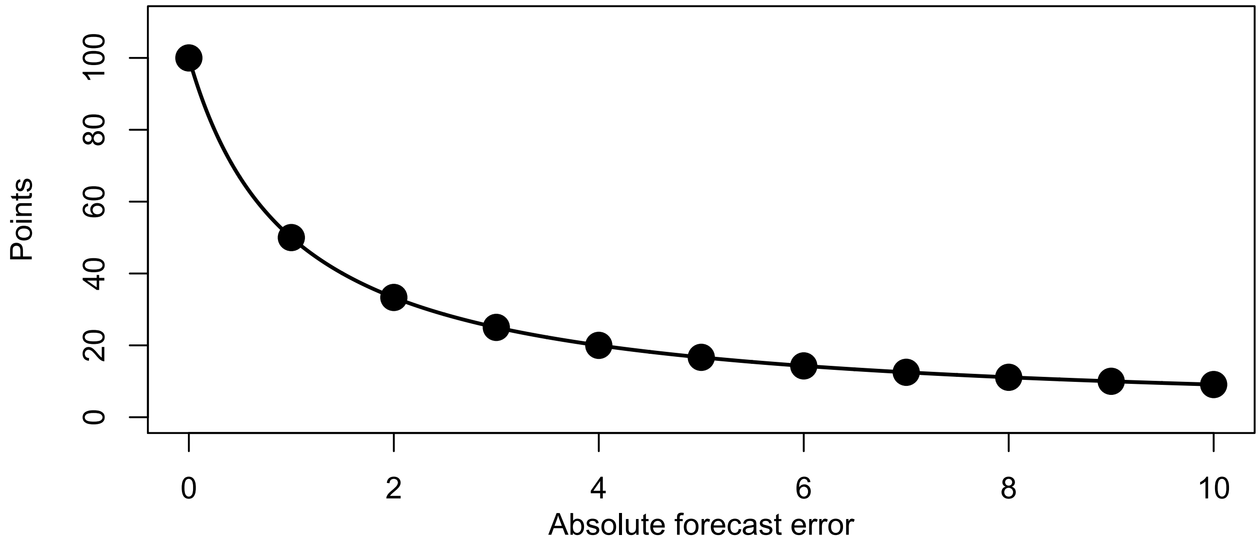


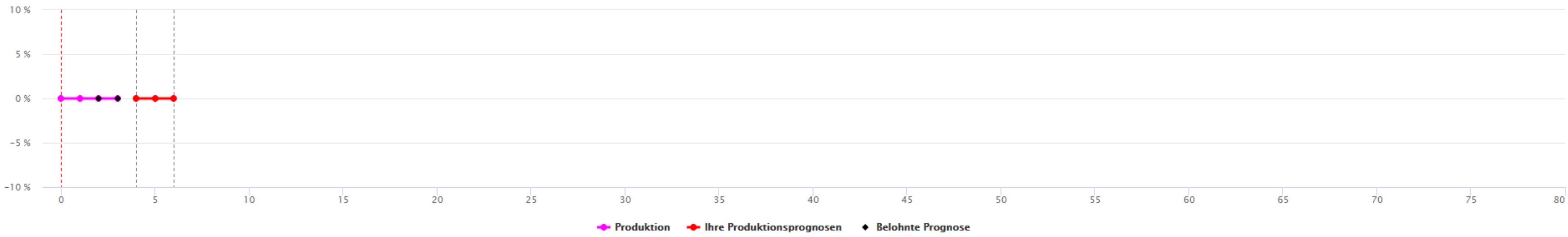
Figure 2: Verhältnis zwischen Ihren Prognosefehlern und den Punkten, die Sie erhalten haben

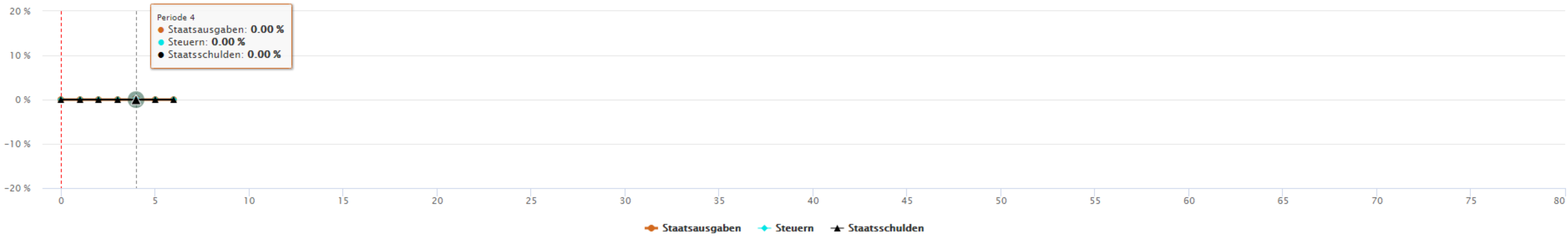
Beispiel: Nehmen Sie an, Sie sind in Periode 11, die Produktion in Periode 10 wird bekannt gegeben und diese beträgt 1,5 %. Sie werden für eine Prognose belohnt, die Sie für Periode 10 abgegeben haben. Insgesamt haben Sie zwei Prognosen für Periode 10 abgegeben: eine in Periode 8 und eine in Periode 9. Jede Prognose wird mit einer Wahrscheinlichkeit von $\frac{1}{2}$ für die Auszahlung herangezogen. Stellen Sie sich vor, dass die in Periode 8 abgegebene Prognose ausgewählt (und dann in der oberen Grafik auf Ihrem Bildschirm angezeigt) wird und Sie eine Produktionsprognose von -4 % abgegeben haben. Da die tatsächliche Produktion 1.5 % beträgt, liegt Ihr absoluter Prognosefehler bei $\mid 1,5-(-4) \mid = 5,5$ Prozentpunkten. Die Abbildung zeigt Ihnen nun an, dass Sie $\frac{100}{1+5,5} = 15$ Punkte für Ihre Prognose erhalten.

Wir befinden uns nun in Periode 4 und Sie müssen Produktionsprognosen für die Perioden 5 bis 6 abgeben.

Aus Ihren Prognosen für Periode 3 , wurde Ihre in Periode 2 abgegebene Prognose zufällig für die Auszahlung ausgewählt. Ihr Prognosefehler betrug 0,0 %. Ihre Punktzahl in dieser Runde beträgt daher 100 Punkte. Ihre kumulierte Auszahlung beträgt aktuell 200 Punkte.

Table 0:35





Ihre Produktionsprognose für die Periode: t=5 t=6

Reset

0% 0%

Wenn Sie mir Ihren Prognosen zufrieden sind, klicken Sie bitte auf „Absenden".

Absenden

Quiz

1. Wie viele Analysten gibt es in der experimentellen Wirtschaft, an der Sie teilnehmen?
2. Angenommen, Sie kommen in Periode 22. Für welche Periode**(n)** müssen Sie eine Prognose abgeben?
3. Wenn Sie in Periode 6 sind, haben Sie Informationen über
   - Die Produktion bis Periode $\cdots$
   - Staatsausgaben, Steuern und Staatsschulden bis Periode $\cdots$
4. Angenommen, die Regierung verringert ihre Staatsausgaben. Welche der folgenden **Punkte** gelten?
   (a) Die Produktion wird mit Sicherheit steigen.
   (b) Die Produktion wird mit Sicherheit sinken.
   (c) Die Auswirkung auf die Produktion hängt von den Auswirkungen auf die durchschnittlichen Produktionsprognosen ab.
   (d) Die Produktion wird wahrscheinlich sinken.
5. Angenommen, Sie befinden sich in Periode 7 und prognostizieren die Produktion in Periode 8.
   (a) Meine Auszahlung in Periode 8 wird definitiv von dieser Prognose abhängen.
   (b) Meine Auszahlung in Periode 8 wird definitiv nicht von dieser Prognose abhängen.
   (c) Meine Auszahlung in Periode 8 hängt von dieser Prognose ab, wenn diese Prognose diejenige ist, die zufällig in Periode 9 ausgewählt und für die Auszahlung herangezogen wird.
   (d) Meine Auszahlung in Periode 8 hängt von dieser Prognose ab, wenn diese Prognose diejenige ist, die zufällig in Periode 8 ausgewählt und für die Auszahlung herangezogen wird.
6. Angenommen, Sie sind in Periode 34. Ihre Prognose für Periode 33, welche Sie in Periode 32 abgegeben haben, wurde zufällig ausgewählt und für die Auszahlung herangezogen. Weiterhin nehmen wir an, dass die tatsächliche Produktion in Periode 33 bei -0.5 % lag und Ihre in Periode 32 abgegebene Prognose bei 0.5 % lag.
   - Wie hoch ist Ihr absoluter Prognosefehler?
   - Wie viele Punkte erhalten Sie in Periode 33?

Wenn Sie alle Fragen beantwortet haben,
BITTE HEBEN SIE IHRE HAND!

Anleitung

Herzlich willkommen zu diesem Experiment! Dieses Experiment ist anonym und die Daten aus Ihren Entscheidungen werden nur mit der Stations-ID verknüpft und nicht mit ihrem Namen. Sie werden am Ende, nachdem alle Teilnehmer das Experiment abgeschlossen haben, bezahlt. Wir behalten uns das Recht vor, Ihre Zahlung zu Ihren Gunsten zu verbessern, wenn die durchschnittlichen Auszahlungen in Ihrer Gruppe niedriger sind als erwartet. Auf Ihrem Tisch liegt ein Taschenrechner, den Sie während des Experiments benutzen können.

Während des Experiments dürfen Sie nicht mit anderen Teilnehmern kommunizieren. Sobald Sie zu irgendeinem Zeitpunkt eine Frage haben sollten, heben Sie bitte die Hand und es wird jemand zu Ihrem Tisch kommen.

Informationen über die fiktive Ökonomie

Das Experiment basiert auf einer Simulation, welche die Schwankungen der realen Ökonomie approximiert. Die Ökonomie, an der Sie teilnehmen, wird hauptsächlich durch **vier Variablen** beschrieben: **Gesamtproduktion, Staatsausgaben, Steuern und Staatsschulden**. Ihre Aufgabe ist es als **professioneller Analyst** zu agieren und Echtzeit-Prognosen über die **zukünftige Produktion** in dieser simulierten Ökonomie zu erstellen. Es gibt neben Ihnen **neun weitere Analysten** in dieser Ökonomie. Die Gruppenzusammensetzung wird sich im Laufe des Experiments nicht verändern. Alle Teilnehmer haben dieselbe Aufgabe.

Die Anleitung wird nun erklären, was Produktion, Staatsausgaben, Steuern und Staatsverschuldung sind, wie sie sich in dieser Ökonomie entwickeln und wie sie von den Prognosen aller Analysten in der Ökonomie abhängen.

Die Werte von Produktion, Staatsausgaben, Steuern und Staatsschulden werden in ***Prozentpunkten*** angegeben, wie es häufig in den Beschreibungen von Volkswirtschaften der Fall ist. Alle Werte können zu jedem Zeitpunkt **positiv**, **negativ** oder **Null** sein und geben lediglich an, ob eine wirtschaftliche Variable **höher**, **niedriger** oder genau **auf ihrem *normalen* bzw. üblichen Niveau ist**. Ein Wert von 1 % bedeutet beispielsweise, dass die Variable 1 % über ihrem normalen Niveau liegt, ein Wert von -8,5 % wiederum bedeutet, dass die Variable 8,5 % unter ihrem normalen Niveau liegt.

Die Produktion ist die Gesamtproduktion von Gütern in der Ökonomie. Die gesamte Produktion wird von den Haushalten (drei Viertel der Gesamtproduktion) und vom Staat (ein Viertel der Gesamtproduktion) konsumiert. Die Produktion ist somit die Summe aus dem Konsum der Haushalte und den Staatsausgaben.

Die Produktion wird im oberen Diagramm auf Ihrem Bildschirm angezeigt (siehe separates Blatt auf Ihrem Schreibtisch). Auf der horizontalen Achse befinden sich die Zeitperioden; die vertikale Achse ist in Prozentpunkten angegeben und zeigt, wie die Produktion (magenta Linie) von ihrem normalen Niveau abweicht. Das normale Niveau beträgt **0 %**. Die Produktion kann jedoch aus zwei Gründen **höher** (d.h. **über 0 %**) oder **niedriger** (d.h. **unter 0 %**) sein.

Der erste Grund ist, dass sich der **Konsum der Haushalte aufgrund Ihrer Prognosen und der Prognosen der anderen Wirtschaftsteilnehmer (Analysten) ändern kann**. Konkret nutzen die Haushalte in der Ökonomie die **durchschnittlichen Produktionsprognosen** der Analysten (auch Ihre Prognosen), um über ihren Konsum zu entscheiden.

Der zweite Grund, warum die Produktion von ihrem normalen Niveau abweichen kann, liegt in der Tatsache, dass sich die Staatsausgaben ändern können. Der Staat finanziert seine Ausgaben entweder durch die sofortige **Erhöhung der Steuern** auf die Haushalte oder durch **staatliche Schuldverschreibungen**. Der Staat hat in der Regel keine Schulden, aber wenn er welche anhäuft, muss er die Schulden **immer zu irgendeinem Zeitpunkt in der Zukunft durch die Erhebung von Steuern auf die Haushalte finanzieren**, d.h. der Staat kann seine Schulden nicht unbegrenzt erhöhen.

Wie also die Staatsausgaben genau finanziert werden, spielt für die Produktion keine direkte Rolle. Ob die Staatsausgaben durch Schulden oder Steuern finanziert werden, kann jedoch die Produktion **beeinträchtigen, wenn es sich auf Ihre Prognosen oder die Prognosen der anderen Analysten auswirkt.**

Die zweite Grafik zeigt die Entwicklung der Staatsverschuldung (schwarze Linie), der Steuern (blaue Linie) und der Staatsausgaben (braune Linie). Vgl. das separate Blatt auf Ihrem Tisch. Auch hier ist die y-Achse in Prozentpunkten angegeben und das normale Niveau jeder dieser Variablen liegt bei 0 %. Die Variablen können aber höher oder niedriger als 0 % werden.

Wenn die Staatsausgaben steigen, **wird die Produktion ebenfalls automatisch ansteigen**. Da sich die Produktion jedoch **nur zum Teil** aus den Staatsausgaben zusammensetzt, würde ein Anstieg der Staatsausgaben die Produktion **nicht in gleichem Maße** beeinflussen.

Darüber hinaus **verringert** ein Anstieg der Staatsausgaben auch **den Konsum**, da die Haushalte entweder sofort mehr Steuern zahlen müssen (wenn die Ausgaben durch sofortige Steuererhöhungen finanziert werden) oder in Zukunft **mehr Steuern zahlen** müssen (wenn die Ausgaben durch Schulden finanziert werden).[1] Dieser negative Effekt auf den Konsum ist kleiner als der automatische Anstieg der Produktion, der infolge einer Staatsausgabenerhöhung eintritt.

Daher ***dürfte die Produktion steigen, wenn die Staatsausgaben steigen***. Der **Gesamteffekt eines Anstiegs der Staatsausgaben auf die Produktion hängt jedoch auch von Ihren Produktionsprognosen und den Prognosen der anderen Analysten** ab. Wie bereits oben erläutert ***führt eine Erhöhung der durchschnittlichen Prognose zu einer Steigerung der Produktion***.

Diese verschiedenen Effekte eines Anstiegs der Staatsausgaben auf die Produktion sind in Abbildung 1 unten zusammengefasst.

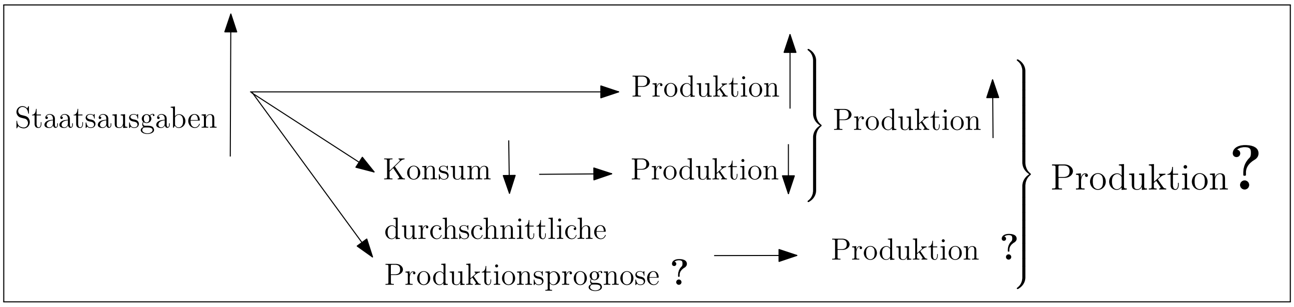


Figure 1: Unterschiedliche Effekte eines Anstiegs der Staatsausgaben auf die Produktion.

Die gleiche Argumentation gilt umgekehrt im Falle eines **Rückgangs der Staatsausgaben**: **Die Produktion dürfte sinken, aber der Gesamteffekt hängt von der durchschnittlichen Produktionsprognose ab** (***ein Rückgang der durchschnittlichen Prognose verringert die Produktion***).

Ihre Aufgabe

Das Experiment dauert 80 Perioden. In jeder Periode müssen Sie **sechs Produktionsprognosen abgeben: eine für jede der nächsten sechs Perioden**. Sie müssen Ihre Prognosen mit den **sechs Reglern**, welche sich unter den Diagrammen befinden, erfassen

[1] Zusätzlich wird eine Erhöhung der Ausgaben der Regierung zu einer Erhöhung des realen Zinssatzes führen, was wiederum die Haushalte dazu veranlasst, mehr zu sparen und weniger zu konsumieren.

(siehe ebenso separates Blatt). Wenn Sie mit Ihren Prognosen zufrieden sind, klicken Sie auf die Schaltfläche „`Absenden`“.

In Periode 1 müssen Sie Produktionsprognosen für die Perioden 2, 3, 4, 5, 6 und 7 abgeben. Sobald jeder Teilnehmer seine/ihre Prognosen abgegeben hat, wird die Produktion in Periode 1 angezeigt und das Experiment geht in Periode 2 über. Anschließend müssen Sie die Produktionsprognosen für die Perioden 3, 4, 5, 6, 7 und 8 abgeben. In Periode 3 müssen Sie die Produktion für die Perioden 4 bis 9 usw. prognostizieren.

Dies bedeutet, dass **Sie insgesamt sechs Produktionsprognosen für den gleichen Zeitraum abgeben müssen**. Beispielsweise müssen Sie die Produktion in Periode 15 zum ersten Mal in Periode 9, dann wieder in Periode 10, 11, 12, 13 und zuletzt in Periode 14 prognostizieren. Im oberen Diagramm zeigen wir neben der Produktion auch Ihre Produktionsprognosen an. Jedoch zeigt die Grafik nicht alle Ihre bisher abgegebenen Prognosen an, sondern nur die jeweils für die Auszahlung herangezogene Prognose für frühere Perioden (schwarze Punkte; sehen Sie unten, wie diese Prognosen ausgewählt werden). Ebenso wird Ihre jüngste Prognose für die aktuelle Periode und die zukünftigen Perioden angezeigt.

Diese Informationen reichen bis zur vorherigen Periode.

Die Regierung wird **immer ihre Staatsausgaben, Steuern und Staatsschulden bekannt geben und diese Ankündigungen umsetzen**. Diese Ankündigungen werden im unteren Diagramm angezeigt. Dadurch haben Sie immer Informationen über die Variablen der entsprechenden Perioden, für die Sie die Produktion prognostizieren müssen.

Auf dem Bildschirm werden in Periode 1 die sechs Regler auf das normale Niveau, nämlich 0 %, initialisiert. In allen anderen Perioden werden die ersten fünf Regler mit den Werten initialisiert, die Sie in der Vorperiode prognostiziert haben. Der sechste Regler, welcher der neuen Periode entspricht, für die Sie noch keine Prognose abgegeben haben, wird mit dem Ausgangswert von 0 % initialisiert. Sie können verschiedene Prognosewerte durch Verschieben des Reglers ausprobieren, wobei die entsprechende Prognose über die Produktionsentwicklung der nächsten sechs Perioden (rote Linie) in der ersten Grafik automatisch angepasst wird. Wenn Sie die Werte wieder auf ihre Initialisierungswerte zurücksetzen möchten, können Sie auf die Schaltfläche „`Reset`“ klicken. Beachten Sie, dass Sie die Diagramme durch Klicken und Ziehen vergrößern können (um den gewünschten Ausschnitt auszuwählen und zu vergrößern).

In der oberen rechten Ecke Ihres Bildschirms wird Ihnen die verbleibende Zeit angezeigt, die Sie für die Eingabe Ihrer Prognosen noch zur Verfügung haben. Wenn die Zeit abgelaufen ist, erscheint die Meldung „Die Zeit ist vorbei!!!“. Sie können Ihre Prognosen auch dann noch abgeben, wenn die Zeit abgelaufen ist, aber in Ihrem eigenen Interesse bitten wir Sie, Ihre Prognosen schnell einzureichen, **um das Experiment nicht zu verzögern**. Zu Beginn des Experiments haben Sie etwas mehr Zeit, um sich mit Ihrer Aufgabe vertraut zu machen.

Zusammenfassend haben Sie in jeder Periode die folgenden Informationen:

- Die Entwicklung der Produktion und Ihre in der Vergangenheit für die Auszahlung herangezogenen Produktionsprognosen (bis zur letzten Periode);
- Die bisherige und zukünftige Entwicklung der Staatsausgaben, Steuern und Staatsschulden (bis zu sechs Perioden in der Zukunft);

All diese Informationen können für die Erstellung Ihrer Prognosen relevant sein, aber es liegt an Ihnen, wie Sie diese verwenden. Sie können diese Informationen auch in einer Tabelle sehen, indem Sie auf die Schaltfläche „`Table`“ oben rechts auf dem Bildschirm klicken. Im späteren Verlauf müssen Sie möglicherweise in der Tabelle nach unten blättern, um alle früheren Perioden anzuzeigen. In der Tabelle sehen Sie auch die Anzahl der Punkte, die Sie bereits gesammelt haben. Wir erklären nun, wie diese berechnet werden.

Informationen zur Auszahlung

Während des Experiments sammeln Sie Punkte. **Die Anzahl der Punkte hängt von der Genauigkeit Ihrer Prognosen ab**. Die Genauigkeit wird durch den absoluten Abstand zwischen dem tatsächlichen Wert der Produktion und einem Ihrer Prognosewerte für diesen Zeitraum gemessen. Diese für die Auszahlung herangezogene Produktionsprognose wird nach dem Zufallsprinzip aus allen sechs Prognosen, die Sie für diesen Zeitraum abgegeben haben, ausgewählt (alle sechs Prognosewerte haben die gleiche Wahrscheinlichkeit ausgewählt zu werden). Die für die Auszahlung herangezogenen Prognosen („belohnten Prognosen“) sind diejenigen, die in der oberen Grafik (schwarze Punkte) dargestellt und in der Tabelle ausgewiesen werden.

Abbildung 2 zeigt die Anzahl der Punkte, die Sie in Abhängigkeit von Ihrem absoluten Prognosefehler erhalten, welcher durch $\frac{100}{1+\text{absoluter Fehler}}$ berechnet wird. Die maximale Anzahl von Punkten pro Periode beträgt 100, welche Sie bei perfekter Prognose erhalten (kein Prognosefehler). Die Tabelle auf Ihrem Bildschirm zeigt die Punkte, die Sie in jeder Periode erhalten haben. Am Ende des Experiments wird Ihr Gesamtpunktestand in Euro umgewandelt und Ihnen anschließend ausgezahlt. **Ein Euro** entspricht hierbei **300 Punkten**.

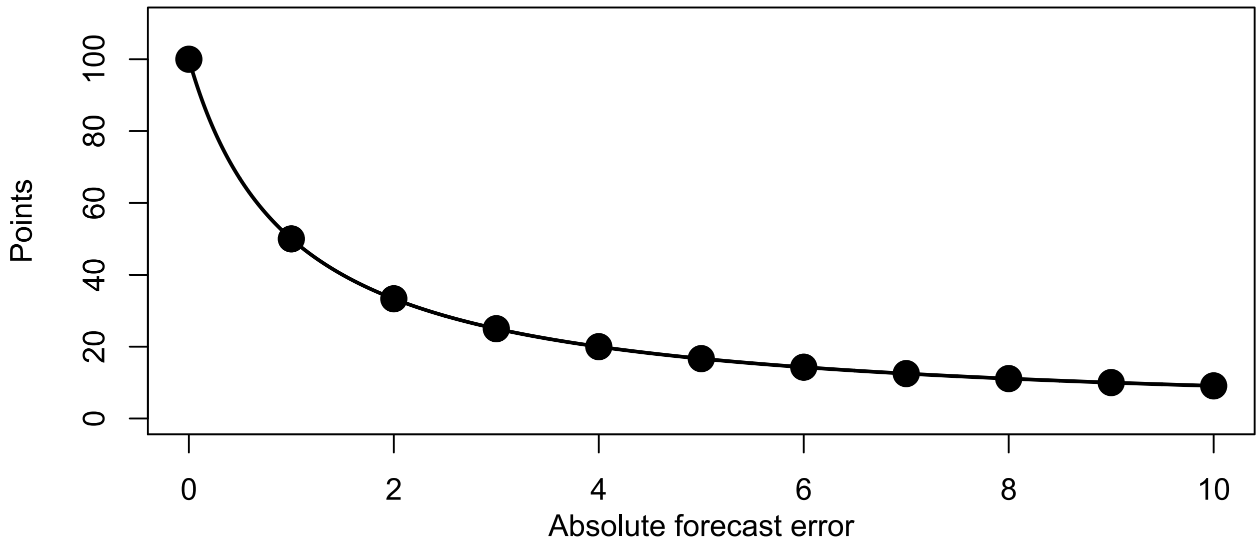


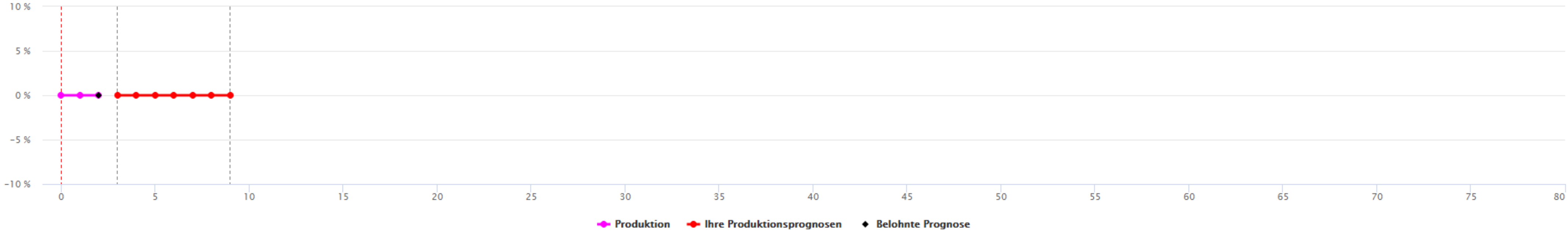
Figure 2: Verhältnis zwischen Ihren Prognosefehlern und den Punkten, die Sie erhalten haben

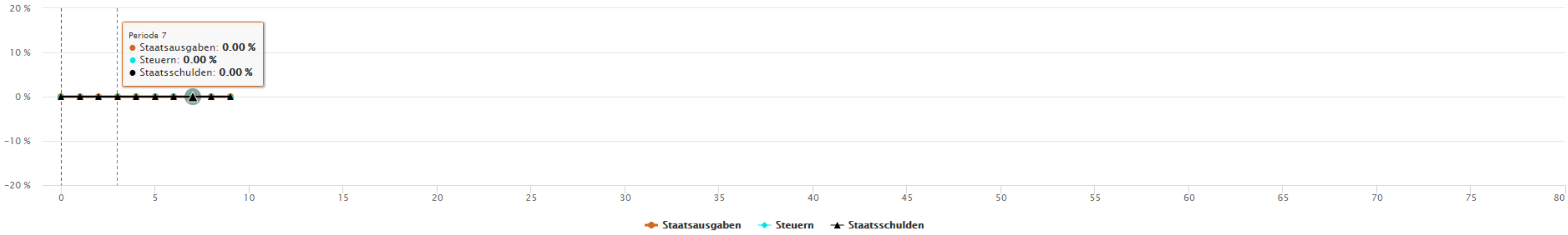
Beispiel: Nehmen Sie an, Sie sind in Periode 11, die Produktion in Periode 10 wird bekannt gegeben und diese beträgt 1,5 %. Sie werden für eine Prognose belohnt, die Sie für Periode 10 abgegeben haben. Insgesamt haben Sie sechs Prognosen für Periode 10 abgegeben: eine in Periode 4, eine in Periode 5, eine in Periode 6, eine Periode 7, eine in Periode 8 und eine in Periode 9. Jede Prognose wird mit einer Wahrscheinlichkeit von $\frac{1}{6}$ für die Auszahlung herangezogen. Stellen Sie sich vor, dass die in Periode 8 abgegebene Prognose ausgewählt (und dann in der oberen Grafik auf Ihrem Bildschirm angezeigt) wird und Sie eine Produktionsprognose von -4 % abgegeben haben. Da die tatsächliche Produktion 1,5 % beträgt, liegt Ihr absoluter Prognosefehler bei $\mid 1,5-(-4) \mid = 5,5$ Prozentpunkten. Die Abbildung zeigt Ihnen nun an, dass Sie $\frac{100}{1+5,5} = 15$ Punkte für Ihre Prognose erhalten.

Wir befinden uns nun in Periode 3 und Sie müssen Produktionsprognosen für die Perioden 4 bis 9 abgeben.

Aus Ihren Prognosen für Periode *2* , wurde Ihre in Periode *1* abgegebene Prognose zufällig für die Auszahlung ausgewählt. Ihr Prognosefehler betrug *0,0 %*. Ihre Punktzahl in dieser Runde beträgt daher *100* Punkte. Ihre kumulierte Auszahlung beträgt aktuell *100* Punkte.

Table 0:53





Ihre Produktionsprognose für die Periode:

Reset

t=4	t=5	t=6	t=7	t=8	t=9
0%	0%	0%	0%	0%	0%

Wenn Sie mir Ihren Prognosen zufrieden sind, klicken Sie bitte auf „Absenden“.

Absenden

Quiz

1. Wie viele Analysten gibt es in der experimentellen Wirtschaft, an der Sie teilnehmen?

2. Angenommen, Sie kommen in Periode 22. Für welche Periode**(n)** müssen Sie eine Prognose abgeben?

3. Wenn Sie in Periode 6 sind, haben Sie Informationen über

   - Die Produktion bis Periode ···
   - Staatsausgaben, Steuern und Staatsschulden bis Periode ···

4. Angenommen, die Regierung verringert ihre Staatsausgaben. Welche der folgenden **Punkte** gelten?

   (a) Die Produktion wird mit Sicherheit steigen.

   (b) Die Produktion wird mit Sicherheit sinken.

   (c) Die Auswirkung auf die Produktion hängt von den Auswirkungen auf die durchschnittlichen Produktionsprognosen ab.

   (d) Die Produktion wird wahrscheinlich sinken.

5. Angenommen, Sie befinden sich in Periode 7 und prognostizieren die Produktion in Periode 10.

   (a) Meine Auszahlung in Periode 10 wird definitiv von dieser Prognose abhängen.

   (b) Meine Auszahlung in Periode 10 wird definitiv nicht von dieser Prognose abhängen.

   (c) Meine Auszahlung in Periode 10 hängt von dieser Prognose ab, wenn diese Prognose diejenige ist, die zufällig in Periode 11 ausgewäht und für die Auszahlung herangezogen wird.

   (d) Meine Auszahlung in Periode 10 hängt von dieser Prognose ab, wenn diese Prognose diejenige ist, die zufällig in Periode 8 ausgewäht und für die Auszahlung herangezogen wird.

6. Angenommen, Sie sind in Periode 34. Ihre Prognose für Periode 33, welche Sie in Periode 30 abgegeben haben, wurde zufällig ausgewählt und für die Auszahlung herangezogen. Weiterhin nehmen wir an, dass die tatsächliche Produktion in Periode 33 bei -0.5 % lag und Ihre in Periode 30 abgegebene Prognose bei 0.5 % lag.

   - Wie hoch ist Ihr absoluter Prognosefehler?
   - Wie viele Punkte erhalten Sie in Periode 33?

Wenn Sie alle Fragen beantwortet haben,
BITTE HEBEN SIE IHRE HAND!

C.1 Post-experiment questionnaire

At the end of the experiment, subjects were asked to answer the following seven questions, with answer keys going from 1 to 5, 1 being ”completely disagree”, 5 being ”totally agree” and 3 being ”neither agree nor disagree”:

Q1 When I made my decision, I thought carefully about how to form my forecast.

Q2 When I saw a change in future government expenditures, I tried to incorporate this into my output forecasts.

Q3 When I saw a change in future government expenditures, I did not incorporate it immediately into my output forecasts, but waited to see whether and how it will affect output.

Q4 I completely ignored changes in government expenditures when submitting my output forecasts.

Q5 I completely ignored changes in government debt when submitting my output forecasts.

Q6 I completely ignored changes in taxes when submitting my output forecasts.

Q7 I found it difficult to think about the experimental worked.

We interpret Q1 as the effort level that subjects put into their experiment task (the higher the score, the higher the level of effort); Q2 as a measurement of immediate reaction to news (the higher the score, the quicker the reaction to news); Q3 as a measurement of delayed reaction to news (a higher score indicates a stronger but delayed reaction to news); Q4 as the level of attention to the spending shocks (the higher the less attention to the spending shocks); Q5 as the level of attention to the government debt (the higher the less attention to the debt); Q6 as the level of attention to the tax level (the higher the less attention to tax) and Q7 as the level of confusion or cognitive load involved (a higher score indicates a higher cognitive load or more confusion).

The participants also had the opportunity to leave us written comments in an open-ended final question.

C.2 GUI when $H = 2$

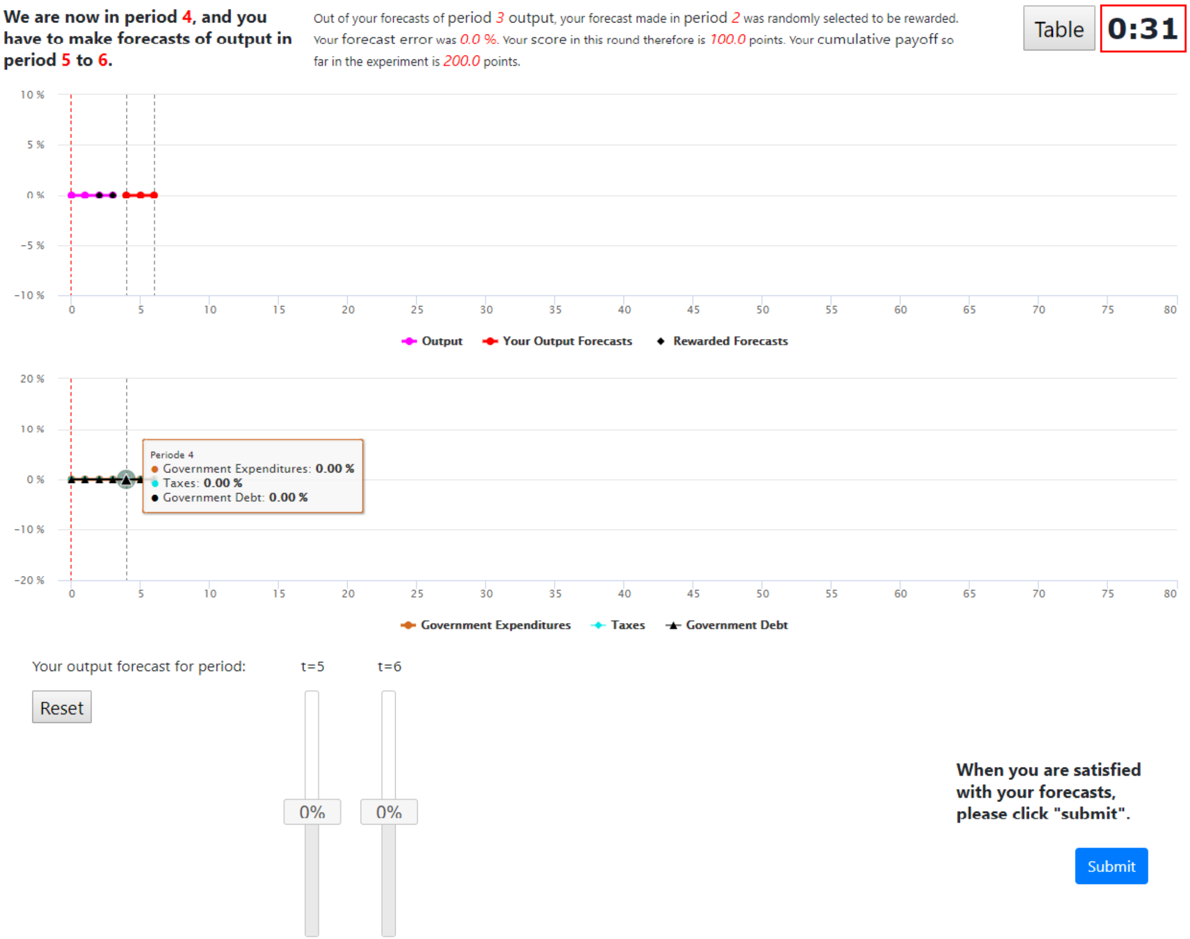


Figure 11: Example of the graphical user interface (GUI) with $H = 2$

C.3 GUI when $H = 6$

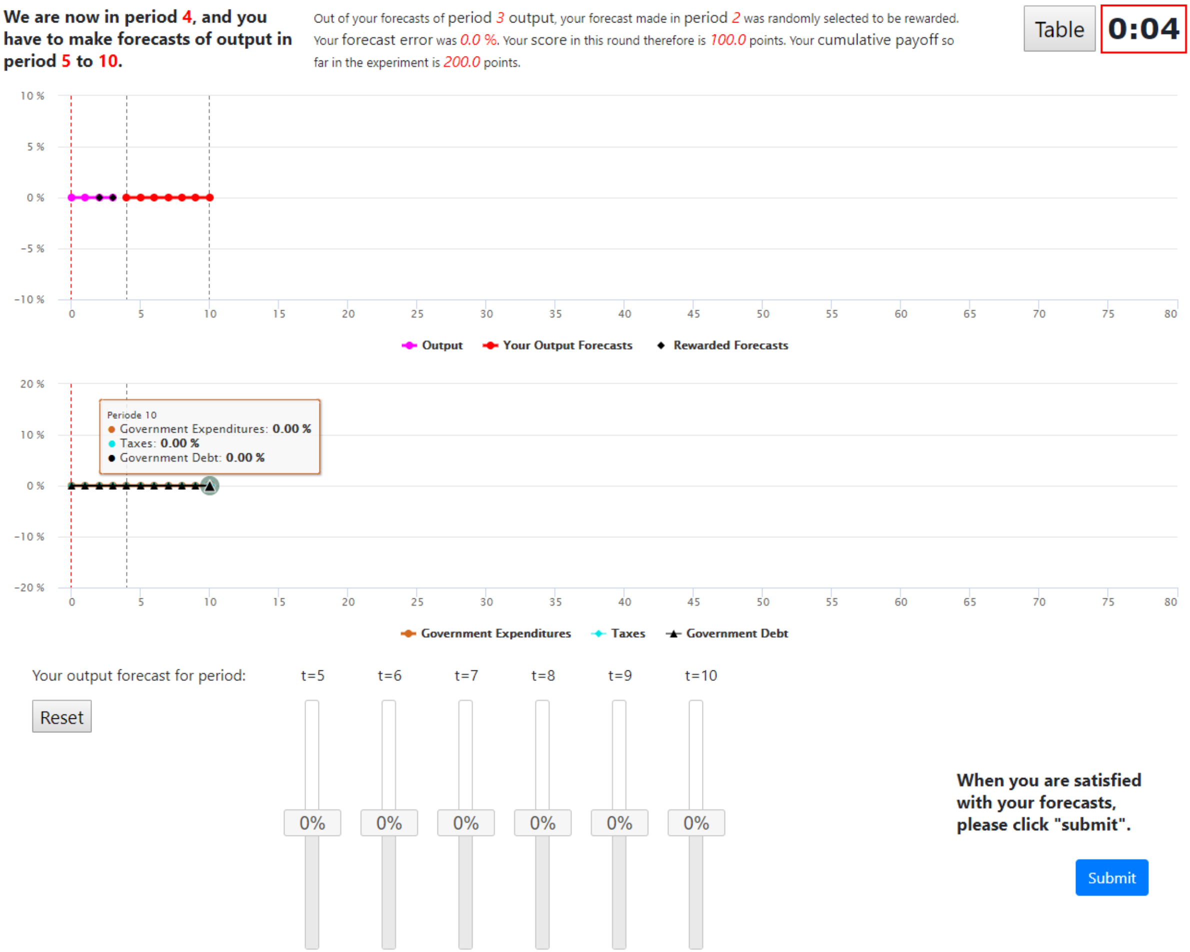


Figure 12: Example of the graphical user interface (GUI) with $H = 6$