ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
*The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics*

Deversi, Marvin

**Working Paper**

# Do Macroeconomic Shocks Affect Intuitive Inflation Forecasting? An Experimental Investigation

Ruhr Economic Papers, No. 528

**Provided in Cooperation with:**
Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)

Suggested Citation: Deversi, Marvin (2014) : Do Macroeconomic Shocks Affect Intuitive Inflation Forecasting? An Experimental Investigation, Ruhr Economic Papers, No. 528, ISBN 978-3-86788-604-8, http://dx.doi.org/10.4419/86788604

This Version is available at:
http://hdl.handle.net/10419/104740

**Nutzungsbedingungen:**
Die ZBW räumt Ihnen als Nutzerin/Nutzer das unentgeltliche, räumlich unbeschränkte und zeitlich auf die Dauer des Schutzrechts beschränkte einfache Recht ein, das ausgewählte Werk im Rahmen der unter
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
nachzulesenden vollständigen Nutzungsbedingungen zu vervielfältigen, mit denen die Nutzerin/der Nutzer sich durch die erste Nutzung einverstanden erklärt.

**Terms of use:**
*The ZBW grants you, the user, the non-exclusive right to use the selected work free of charge, territorially unrestricted and within the time limit of the term of the property rights according to the terms specified at*
→ *http://www.econstor.eu/dspace/Nutzungsbedingungen*
*By the first use of the selected work the user agrees and declares to comply with these terms of use.*

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft

RUHR
ECONOMIC PAPERS

Marvin Deversi

# Do Macroeconomic Shocks Affect Intuitive Inflation Forecasting?

## An Experimental Investigation

RUHR UNIVERSITÄT BOCHUM RUB

#528

# Imprint


## Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany



## Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, e-mail: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Roland Döhrn, Prof. Dr. Manuel Frondel, Prof. Dr. Jochen Kluve
RWI, Phone: +49 (0) 201/81 49-213, e-mail: presse@rwi-essen.de

## Editorial Office

Sabine Weiler
RWI, Phone: +49 (0) 201/81 49-213, e-mail: sabine.weiler@rwi-essen.de



## Ruhr Economic Papers #528

Responsible Editor: Thomas Bauer



All rights reserved. Bochum, Dortmund, Duisburg, Essen, Germany, 2014



ISSN 1864-4872 (online) – ISBN 978-3-86788-604-8



The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions and do not necessarily reflect those of the editors.

Ruhr Economic Papers #528

Marvin Deversi

# Do Macroeconomic Shocks Affect Intuitive Inflation Forecasting?

## An Experimental Investigation

RUHR UNIVERSITÄT BOCHUM RUB

## Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788604
ISSN 1864-4872 (online)
ISBN 978-3-86788-604-8

Marvin Deversi[1]

# Do Macroeconomic Shocks Affect Intuitive Inflation Forecasting?

## An Experimental Investigation

## Abstract

*In an experimental setting impulse-response behaviour in intuitive inflation forecasting is analysed. Participants were asked to forecast future values of inflation for a fictitious economy after receiving charts and lists of past values of inflation and output gap. Thirty periods were forecasted stepwise and feedback on performance was provided after each period. In a between subjects design, participants experienced a negative or positive supply shock. The results suggest that participants barely report rational forecasts. Instead, simple backward-looking rules describe stated forecast series. Forecasting is heterogeneous across agents and over time. Before the shock, most participants can be described by natural expectations. Due to the shocks 69% of participants are found to switch their forecasting rule. After the negative supply shock, subjects increase efficiency of forecasts. But, after a positive supply shock efficiency drops down to zero; this is evidence for a negativity bias. As a main result, macroeconomic shocks do alter the way experimental participants form intuitive inflation forecasts, however, to what extent depends on the shocks' characteristics.*

*JEL Classification: C91, D84, E03*

*Keywords: macroeconomic experiment; inflation expectations; intuitive forecasting; shocks; heterogeneity*

*November 2014*

*1 Marvin Deversi, University of East Anglia and Ruhr University Bochum. – All correspondence to Marvin Deversi, Institute for Macroeconomics & Bochum Lab for Experimental Economics (RUBex), Faculty of Economics, Ruhr-University Bochum, Universitätsstr. 150, 44801 Bochum, Germany, e-mail: marvin.deversi@rub.de*

> "When the crisis came, the serious limitations of existing economic and financial models immediately became apparent. Macro models [...] seemed incapable of explaining what was happening to the economy in a convincing manner. [*To improve this*] we may need to consider a richer characterisation of expectation formation. Rational expectations theory has brought macroeconomic analysis a long way over the past four decades. But there is a clear need to re-examine this assumption."
>
> (Trichet, 2010, November 18)

# 1 Introduction

Monetary policy makers need information about how people form their inflation expectations to successfully create policy. In his opening address at the ECB Central Banking Conference, Jean-Claude Trichet, former president of the European Central Bank, criticises that standard macroeconomics has not delivered this information. And, in fact when large shocks hit an economy, it seems that little is known about inflation expectation formation (Coibion and Gorodnichenko, 2012).

Most macroeconomists advocate the Rational Expectation Hypothesis (REH) as established by Muth (1961) and Lucas (1972). In doing so, economic agents are assumed to homogeneously exhibit model consistent expectations and use all information available. As Lovell (1986) explains, this results in a zero expected forecast error. In the spirit of Trichet's critique, a strand of economic literature contrasts that expectations on inflation and on other macroeconomic variables (e.g. taxes, Bernasconi et al., 2009; unemployment, Garz, 2013; wage levels, Roos and Luhan, 2013; or growth rates, Bovi, 2014) appear to be not perfectly rational. As a general result economic agents are found to be *bounded rational* instead (Hommes, 2011, p. 2). That is, they use subjective mental models to form forecasts about future values of macroeconomic variables, so-called *intuitive forecasts* (Beshears et al., 2013). These forecasts can be seen as numerical representations of peoples' expectations (Roos and Luhan, 2013).[1] Here, several survey-based and experimental studies find that peoples' intuitive forecasting is neither homogeneous across agents nor stable over time (*between* and *within heterogeneity*). In particular, this is true when investigating inflation forecasting. From recent literature it also seems that structural breaks, as caused by unexpected policy changes, affect heterogeneity in intuitive forecasting (Pfajfar and Zakelj, 2011; Odria and Rodriguez, 2013).

[1] Burke and Manz (2011) allude that reporting numerical forecasts is scarce in real life situations but at least done implicitly. Despite there being possible differences, forecasts are used as synonyms for expectations henceforth.

The present study contributes to this literature by experimentally investigating whether macroeconomic shocks affect heterogeneity in intuitive inflation forecasting. Using a learning-to-forecast experiment, panel data is generated to estimate which statistical rules describe participants' individual forecasting best. Then the robustness of the arising forecasting taxonomy is tested in the presence of different shocks.

Methodologically, participants are faced with past values of simulated data from a stochastic aggregate demand and supply (AD-AS) business cycle model. In each of the 30 periods, the participants' task was to forecast the future values of inflation for a fictitious economy one period ahead. After every forecasting period subjects were provided with performance feedback. This information allows them to adjust forecasting over time (*learning to forecast*). Predetermined shocks simulate structural breaks in the inflation trajectory. Treatments vary in the direction of the shock hitting the economy. Therefore, participants experience a negative or positive supply shock.

In the literature, experiments are frequently used to elicit inflation expectations (e.g. Adam, 2007; Bernasconi et al., 2009; Odria and Rodriguez, 2013; Roos and Luhan, 2013; Pfajfar and Zakelj, 2014). This method has some strengths, especially when compared to survey analysis of inflation expectations. In the field, policy changes or other shocks trigger variations in several variables such that respective effects on forecasting can hardly be isolated (Pfajfar and Zakelj, 2014). At the same time expectation reports are influenced by neighbours, the media or experts' forecasts (Carroll, 2003) and refer to a variety of socio-economic conditions (Burke and Manz, 2011). Besides, survey responses are not incentivised (Garz, 2013). Studying intuitive forecasting in a laboratory setting solves these drawbacks. Participants are payed according to their forecasting performance. Then researchers can control for possible confounds by varying the information presented, using artificial data and/or isolating participants from each other (Bernasconi et al., 2009). Also, by deliberately designing a laboratory environment, researchers can test economic models in their own domains. And most importantly, experiments allow clean comparative-static analyses which makes them particularly informative when investigating shock responses of individual behaviour.

The reminder of this study is structured as follows: Firstly, the related literature describing the study's contribution is reviewed. Secondly, used forecasting models are described and the experimental design is explained. Then the results are presented. Section 5 provides a discussion of the findings and reaches an answer towards the proposition in the heading. Section 6 concludes this study.

# 2 Related Literature

In economic literature macroeconomic shocks are attributed as "traumatic events" (Necker and Ziegelmeyer, 2014, p. 1) and found to have lasting effects on individual economic behaviour. Empirical evidence suggests that financial market participants are more risk averse after recessions and respective loss experiences (Malmendier and Nagel, 2011; Guiso et al., 2013). Giuliano and Spilimbergo (2014) use self-reported answers from the General Social Survey and match these with macroeconomic time series. They detect a relationship between macroeconomic shocks and responders' beliefs about the future. For example, individuals who experienced a recession in their early adulthood tend to emphasise the influence of luck on success more strongly. Using evidence from rural Indonesia, Cameron and Shah (2013) state that people who have experienced natural disasters (like earthquakes) expect the probability of a future disaster to be higher compared to people who experienced no disaster.

Similarly, intuitive forecasting in a macroeconomic context seems to be affected by a shock. Bovi (2014) even defines shocks "[...] as unexpected events leading agents to revise their expectations" (p. 2). Using Italian data from the Business Surveys Unit of the European Comission, Bovi (2014) explores how shocks shape the time series properties of reported survey forecasts about the future Gross Domestic Product (GDP). The occurrence of negative shocks is displayed as significant troughs in the Italian GDP. These are found to have persistent effects on the level of agents' expectations. In detail, agents overreact to a contraction in GDP. The degree of forecasting disagreement across responders remains persistently high. Before and after negative shocks, the volatility of stated forecast remains the same.

In contrast, Gnan et al. (2010) document increasing disagreement among individuals after the macroeconomic crisis 2008/2009 for the Euro Area. After the crisis, forecasts' volatility decreased. Gnan et al. (2010) interpret this as a decrease in subjective uncertainty. They explain that the agents' awareness and also the effort they put into forecasting increase after a shock, which in turn reduces uncertainty. Furthermore, the spikes in several time series may induce saliency of the relationships among macroeconomic variables helping to build natural forecasts in future times.

Roos and Luhan (2013) test for several statistical models describing individual inflation forecasts in one experimental environment and report considerably high cross-sectional heterogeneity. In their experiment, participants are assigned to the role of either being a firm or a worker. In each period the worker chooses a nominal wage and the firm its respective amount of labour being employed. Both roles are part of a model macroeconomy and participants' decisions and expectations influence the general price development and

output of the economy.[2] All participants are informed about the underlying model and the core relationships between the model's variables (almost complete information). During the experiment firms and workers are asked to state their forecasts about the next period's general price level. It turned out that about 24% of the participants can be described best by using adaptive expectation formation. Around 19% by using an anchoring-and-adjustment forecasting rule, 16% build natural expectations, 5% a trend extrapolation and 15% use static expectation formation. The rest is either rational or it is not possible to fit the data to one of the presented models. This is representative for what other experimental studies found (see Pfajfar and Zakelj, 2011, 2014; Odra and Rodriguez, 2013). However, Roos and Luhan (2013) do not exogenously induce impulses within ther laboratory macroeconomy.

Pfajfar and Zakelj (2014) do so. The authors explore intuitive inflation forecasting by facing participants with simulated data from a New Keynesian economy. Subjects are given the role of a statistical bureau stating forecasts about future inflation. Also, they are informed that (among other factors) their forecasts and other participants' forecasts influence the real trajectory of future inflation (with market feedback). In particular, Pfajfar and Zakelj (2011)[3] analyse how exogenous variations of institutional features affect individual inflation forecasting. These exogenous variations can be seen as unexpected shocks affecting the trajectories of the time series presented to the participants. Pfajfar and Zakelj vary the conducts of monetary policy and find that individual forecasting rule heterogeneity increases with the passivity of the central bank.[4] Thus, variations in the data generating model seem to influence intuitive forecasting and heterogeneity across participants.

Odria and Rodriguez (2013) also analyse how exogenous unexpected events shape intuitive inflation forecasting, here in an open New-Keynesian macroeconomy. Similar to Pfajfar and Zakelj (2011), participants experience the adoption of a new monetary policy rule (changes in parameters) in one treatment. In two other treatments the effect of publicly announcing this adoption is analysed. Interestingly, the economy is hit by a large negative demand shock midway through the forecasting periods in a fourth treatment. The authors confirm earlier literature by finding that participants are rarely forecasting rationally, but instead use time series extrapolation. The more aggressive the central

[2]This link between individual and aggregate behaviour is known as *market feedback* and a frequently used feature in learning-to-forecast experiments. Next to obtaining clean data on intuitive forecasting, market feedback allows the observation of respective aggregate market outcomes. Since this paper aims to analyse individual forecasting only, I focus on the former.

[3]This study is the companion paper of Pfajfar and Zakelj (2014); however, the former focuses on exogenous variations in the policy conduct, the latter on a new consistency-based rationality test in experiments with market feedback.

[4]Similarly, Assenza et al. (2013) analyse different conducts of monetary policy in an experimental New Keynesian economy with market feedback, but focus on aggregate rather than individual outcomes.

bank aims to stabilise the inflation the more rational forecasts are stated. The negative demand shock is found to decrease the number of participants which are described best by the REH. Despite the shock, subjective uncertainty reports remain persistently high.

Nevertheless, it can be argued that previous experimental approaches to analyse impulse-response beaviour in individual forecasting like in Odria and Rodriguez (2013) cannot cleanly isolate respective effects. In Pfajfar and Zakelj (2011) and Odria and Rodriguez (2013) individual forecast series before and after the exogenous intervention are not comparable in their time series properties. For example, after a shock hits a New Keynesian model the fundamental movements of the core variables change because of the model's persistence. Rather than being a stationary process as is the case before the shock, the movements can be better described by a smooth adjustment back to the pre-phase expected value. Other than that, most learning-to-forecast experiments are featured with a positive market feedback mechanism. Because participants are aware of this mechanism, they start thinking about the shock reaction of the other 11 subjects in the economy (each economy is influenced by 12 players decisions).[5] Thus, the influences on the series become even more complex after the shock. This could result in the variance increasing strongly or in the occurence of pathologic trajectories. Consequently, the observed changes in intuitive forecasting could be due to changes of the time series properties of inflation and not only because participants experienced a negative demand shock or a more aggressive central bank. In line with this duct, these designs are suitable to analyse expectation induced shock responses displayed in macro time series, however, it is suboptimal to analyse effects of experiencing a shock on individual forecasting.

The experiment here takes these drawbacks into account and expands the analysis further. The simulated data is constructed such that the pre- and post-phase forecasting are comparable. By doing so the inflation trajectory after a shock hit the economy is replaced by a model simulation without a shock. And, the forecasting task is not featured with a market feedback. So, differences between forecasting before and after the shock are due to the fact that participants experienced a one-period shock before post-forecasting.

Finally, the main contribution of this paper exists in the comparison between positive and negative shocks. Economic and psychological literature finds evidence for individuals reacting stronger to a negative than to a positive framing of exogenous shocks of comparable magnitude, the so-called *negativity bias* (Holbrook et al., 2001; Soroka, 2006). Garz (2013) confirms a negativity bias when people react to news shocks. Similarly, Lamla

[5]That participants start thinking about how other players react to a shock seems to be unrealistic. O'Connor (2011) state evidence that during and after the crisis in 2008/2009 people report a strong feeling of fatalism. Participants felt that they had no ability to influence the actual economic developments. In Odria and Rodriguez (2013) the degree of feeling fatalism can be assumed to be much lower since the economy consists of 12 participants only.

and Lein (2014) report that information about increasing inflation has a much stronger effect on inflation perception than information about decreasing inflation. Necker and Ziegelmeyer (2014) argue that in times of economic crisis induced emotions are much stronger than in boom times. Soroka (2006) relies on the prospect theory of choice under uncertainty by Kahneman and Tversky (1974) to explain the negativity bias phenomenon. Here, economic agents are assumed to have subjective reference points they relate their behaviour to. Negativity bias is observable if agents assume this reference point to be lower bound such that things cannot get worse. In terms of Benartzi and Thaler (1995), negative scenarios may induce stronger *loss aversion*. Whether an asymmetry between positive and negative shocks is present in intuitive forecasting can be analysed by varying the direction of shocks.

# 3 Methods

## 3.1 Considered Forecasting Models

Lovell (1986) shows that from implications of the REH it follows that peoples' forecasts are unbiased and efficient. Using a regression approach, we can write that the forecast of individual $k$ at period $t$ about inflation ($\pi$) one period ahead ($t+1$) is

$$\pi^k_{t+1|t} = \alpha_0 + \alpha_1 \pi_{t+1} + \epsilon_t, \tag{1}$$

where $\epsilon_t$ is a normally distributed error term. If $\hat{\alpha_0} = 0$ and $\hat{\alpha_1} = 1$ we cannot reject the hypothesis that participants report unbiased forecasts. Mankiw et al. (2004) explain that efficiency can be tested by regressing the individual $k'$s forecast error ($f^k_t$), i.e. the deviation from the stated forecast to the real inflation value, on all past values of output gap. If none of the lags is significant in explaining $f^k_t$, forecasts are deemed to be efficient. This is because rational agents are assumed to exploit all available information to forecast. Elicited inflation expectations being unbiased and efficient are considered to be rational.[6]

From a review of studies investigating macroeconomic expectations via learning-to-forecast experiment, Hommes (2011) summarises that "[...] agents do not know the true law of motion of the economy, but instead use time series observations to form expectations [...]" (p. 2). This backward-looking behaviour can frequently be described by using simple statistical rules (Roos and Luhan, 2013).[7] Admittedly, there does not exist a single rule

[6]Note that the use of econometric rationality tests is not without scepticism. Andolfatto et al. (2008) report that rationality tests reject the REH in a large percentage of cases even though the analysed data was generated with a stochastic macro model incorporating rational expectation formation.

[7]The literature verbally describes these rules in terms of agents literally acting like a rule suggests. Fuster et al. (2010) explain that one should rather think in terms of an 'as if' interpretation. Intuitive

describing peoples' intuitive inflation forecasting best neither when using experimental or survey data (e.g. Branch, 2004; Odria and Rodriguez, 2013). Instead, the *heterogeneity in intuitive inflation forecasting* is a well established fact in the literature (Hommes, 2011). Different survey responders or experiment participants use different rules. Adding strength to this result, recent empirical studies test several backward-looking rules of how people intuitively forecast inflation.

Pfajfar and Zakelj (2014) test a standard descriptive model of expectation formation, *autoregressive expectations* (AE). The idea is that participants use observed past values to form a forecast for the next period. AE can be described as univariate time series models of AR(1), AR(2) or generally AR(p)-form.

$$\pi^k_{t+1|t} = \lambda_1 \pi_{t-1} + \epsilon_t \tag{2}$$

$$\pi^k_{t+1|t} = \lambda_1 \pi_{t-1} + \lambda_2 \pi_{t-2} + \epsilon_t \tag{3}$$

$$\pi^k_{t+1|t} = \sum_{p=1}^{T} \lambda_p \pi_{t-p} + \epsilon_t \tag{4}$$

The *AR*-form seems to be compelling because of its simplicity (Beshears et al., 2013). Model (2) can be referred to as *static* or *naive expectations* (SE) if $\lambda_1 = 1$. Here, agents use the observed inflation in the last period as a forecast for the next period. So,

$$\pi^k_{t+1|t} = \pi_{t-1} + \epsilon_t. \tag{5}$$

Also *trend extrapolation* (TE) is found to be capable describing individual inflation expectation formation (e.g. Pfajfar and Zakelj, 2014; or Roos and Schmidt, 2011). Whereby trend extrapolation is specified as

$$\pi^k_{t+1|t} = \pi_{t-1} + \lambda_T (x_{t-1} - x_{t-2}) + \epsilon_t, \tag{6}$$

so forecasts are formed as a combination of the past value of inflation and the deviation of the past value to its former value (*trend*).

To distinguish these simple extrapolation rules, Evans and Honkapohja (2001) suggest that economic agents behave more like econometricians. They learn by their former forecast errors and adjust forecast rule parameters until forming stable (equilibrium) expectations. This idea is well known as *adaptive learning* (AL). However, unlike for example Least Square Learning or Bayesian Learning theories suggest, people are found to have weak econometric competencies (Roos and Luhan, 2013). The adaptive expectation

forecasting rules do not explain how agents calculate their forecasts but the forecasts can be statistically described as if they would do so.

formula can be interpreted as the simplest form of adaptive learning. Here, people adjust their forecasts to their former forecast errors. As Hey (1994) formalises

$$\pi^k_{t+1|t} = \pi^k_{t|t-1} + \lambda_A(\pi_{t-1} - \pi^k_{t|t-1}) + \epsilon_t, \tag{7}$$

where $\lambda_A$ is an adjustment constant. Among others, Odria and Rodriguez (2013) find that a small proportion of subjects' inflation forecasting (9%) can be described this way. Generally, AL in its many different types is found to be capable of describing a high fraction of peoples' intuitive forecasting in a macroeconomic context (Bernasconi et al., 2009).[8]

Fuster et al. (2010) propose that agents form *natural expectations*. This approach can be described as a mixture of backward-looking and rational expectations. Agents are assumed to combine information of several variables to forecast. They intuitively assess a theoretical relationship between say inflation ($\pi$) and output ($y$) to gain information about the future development of inflation; expectations are theory-driven. In a bivariate model setting, to forecast $\pi_{t+1}$ agents refer to past values of $\pi$ and values of $y$, such that

$$\pi^k_{t+1|t} = \lambda_N \sum_{l=1}^{n} \pi_{t-l} + (1 - \lambda_N) \sum_{j=1}^{m} y_{t-j} + \epsilon_t, \tag{8}$$

where $n$ and $m$ are the numbers of maximal available lags of $\pi$ and $y$, respectively, and $\lambda_N$ is the weight of $\pi$'s past values compared to $y$'s past values.

Models (2)-(8) illustrate that agents seem to reduce complexity of a forecasting task by using simplifying rules instead of knowing the underlying model. For this reason the rules presented here are often referred to as *heuristics* (Hommes, 2011). Gnan et al. (2010) explain that agents are prone to several psychological biases when forecasting future inflation. The authors review heuristics incorporating such biases, whereby one of these rules appears compelling to several studies, the *anchoring and adjustment heuristic* (AA) (e.g. Hey, 1996; Roos and Luhan, 2013). This rule is related to Kahneman and Tversky (1974) stating that agents behaviour in an uncertain environment is based on a subjective reference point (*anchor*). Following Anufriev and Hommes (2012)

$$\pi^k_{t+1|t} = \frac{\frac{\sum_{h=1}^{v} \pi_{t-h}}{v} + \pi_{t-1}}{2} + (\pi_{t-1} - \pi_{t-2}) + \epsilon_t. \tag{9}$$

The AA is a mixture consisting of an extrapolation of the past price change $(\pi_{t-1} - \pi_{t-2})$ and an average expectation of the last observed value of $\pi$ and a long-run trend estimate.

[8] Adaptive learning theories and the idea of stepwise coefficient updating can be based on many simple forecasting rules (e.g. on trend extrapolation or other recursive rules). For an overview see Evans and Honkapohja (2001) or Pfajfar and Santoro (2010).

Anufriev and Hommes (2012) assume this estimate to be a sample average of all observed past values (so $v$ values). $h$ is the number of inflation lags. Individual $k$'s forecast is anchored on the long-run trend approximation $(\frac{\frac{\sum_{h=1}^{v} \pi_{t-h}}{v} + \pi_{t-1}}{2})$ from which a recent trend is continued. The idea that inflation expectations are anchored is emphasised by Malmendier and Nagel (2013), also.

Brock and Hommes (1997) introduce a model of *switching among forecasting rules*. Here, participants tend to switch among a set of forecasting rules according to which of these performed better in the past. For this type of behaviour will be tested, as well.

The degree of heterogeneity in intuitive inflation forecasting is immense. This "wilderness of bounded rationality" (Hommes, 2011, p. 12) has to be disciplined by relying on empirical analyses reporting evidence for possible model formulations. Accordingly, this study focuses on models (1)-(9) to test which rules describe participants' intuitive forecasting best.

## 3.2 Data to be Forecast

The time series being forecasted by experimental participants is constructed using model simulations of a standard closed AD-AS model. The main motivation to use artificially generated data is that the opportunity to refer to information from outside the lab is ruled out. All information can be controlled. Formally, the model can be described by the aggregate demand curve (eq. (10)), the short-run aggregated supply curve (eq. (11)) and the assumption about expectation formation of the representative household (eq. (12)).

$$y_t - \bar{y} = \alpha(\pi^* - \pi_t) + z_t \tag{10}$$

$$\pi_t = \pi_t^e + \gamma(y_t - \bar{y}) + s_t \tag{11}$$

$$\pi_t^e = \phi\pi_{t-1}^e + (1-\phi)\pi_{t-1} \tag{12}$$

Where $y_t - \bar{y}$ is the deviation of output from its potential level (output gap), $\pi^* - \pi_t$ is the deviation of the current inflation rate from its announced target value and $z_t$ is an i.i.d. demand shock variable for which holds that $z_t \sim N(0, 0.81)$. If $z_t > 0$ a positive demand shock occurs, if $z_t < 0$ a negative demand shock takes place. (10) represents the simultaneous equilibrium on the money and the good's market. Whereas (11) describes the supply side, so the price- and wage-setting equilibrium. Here, $\pi_t^e$ is the expectation of the value of inflation in the next period as assumed by the model. (12) shows that this is adaptive inflation expectation formation. $\phi$ is a parameter determining the inertia in expectation formation. The higher $\phi$ the more reluctant the representative household

is to adjust its expectation. Similar to $z_t$, $s_t$ is an i.i.d. supply shock variable for which holds that $s_t \sim N(0, 0.04)$. If $s_t < 0$ a positive supply shock, if $s_t > 0$ a negative supply shock takes place.

This model is simulated twice, once without and once with a large shock taking place in period 34. Between two treatments the direction of the shock varies. The calibration of the model refers to Sorensen and Whitta-Jacobsen (2010). The variables' values are set such that $\phi = 0.9$, $\gamma = 0.075$, $\alpha = 14.4$ and $\pi^* = 0$. In the first treatment ($T^-$) $s_{34} = 1.75$, so a negative supply shock hits the economy. Whereas in the second treatment ($T^+$) a positive supply shock takes place, such that $s_{34} = -1.75$. The model response in the shock period is used to replace the time series values in the model without a shock. Thus, forecasting before and after a shock takes place is comparable. All differences in inflation's expected value, amplitude and variance are due to randomness only. Without this construction rule, both supply shocks would persist in the time series resulting in varying properties of inflation. The simulated data for both treatments is presented in appendix A. Overall, this can be described as a 2x2 factorial design. The first dimension is pre- and post-phase forecasting which is established as a within subject variation. The second dimension is between subjects and varies in the direction of the supply shock taking place.

## 3.3 Experimental Procedure

In a learning-to-forecast experiment each subject is introduced to its role as an agent in a fictitious economy. Participants work in a firm for which they predict future values of inflation.[9] Whereby inflation describes the general price level in the fictitious economy. Also, information about the output gap is provided. It is public knowledge that all rules and the economy itself remain the same in all periods and that stated forecast do not influence the real development of inflation. Participants know that unpredictable events can influence both macroeconomic variables such that these fluctuate around. In the laboratory participants are separated from each other by partitions. The experimenter ensured that there was no talking during the experiment. The instructions were recorded in advance and played aloud before the experiment started.[10] These carefully explain the concepts of inflation and output gap. To ensure that every subject understands the instructions, there was a comprehension task. The experimenter checked all answers and provided verbal help in case one of the questions was answered incorrectly. The experiment

[9] This forecasting context is adapted from Pfajfar and Zakelj (2011, 2014) and Assenza et al. (2013) to ensure comparability. Note, Burke and Manz (2011) explain that forecasting behaviour is not consistent across different contexts. So, findings from context-free settings may not apply to inflation forecasting.

[10] All experimental material can be found in the appendix B.

Table 1: Payoff Function

| $f_t^k$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.9 | 2 |
|---|---|---|---|---|---|---|---|
| $score_t^k$ | 100 | 50 | $\frac{33}{3}$ | 25 | 20 | 10 | 0 |

was run using z-Tree (Fischbacher, 2007).[11] In 32 periods subjects state an inflation forecast one period ahead. For each forecast subjects had 80 seconds.[12] Admittedly, the first two periods are trial periods. These trial periods are intended to make participants familiar with the computer environment and the underlying model economy. Forecasts in the trial periods earn no money.

In addition to £2 for showing up on time, participants earn money according to their forecasting performance. Here, rewards after the experiment has ended increase with the accuracy of their forecasts in every period. The accuracy is measured by the *absolute forecast error* ($f_t^k$), i.e. the deviation between $k$'s stated inflation forecast and the real inflation rate in period $t$.

$$f_t^k = |\pi_t - \pi_{t|t-1}^k|. \tag{13}$$

The higher the distance between forecast and real value, the higher is $f_t^k$ and the lower is the *prediction score* ($score_t^k$) of subject $k$ in period $t$. The prediction score is measured in experimental points and reflects the forecasting accuracy via the forecast error. However, if $f_t^k > 2$ the achieved score reduces to 0. Such that

$$score_t^k = \begin{cases} \frac{10}{0.1+f_t^k} & \text{if } f_t^k < 2 \\ 0 & \text{if } f_t^k \geq 2 \end{cases}. \tag{14}$$

This payoff function incentivises participants to forecast future inflation most accurately.[13] Table 1 presents prediction scores for selected values of $f_t^k$. The total prediction score of all 30 periods in experimental points is transferred into pounds at an exchange rate of 0.012, so that 100 experimental points are worth £1.20. Perfectly predicting the whole series yields £36.

[11] An example forecasting screen is presented in appendix A.

[12] The time constraint aims to skimp the duration of the whole experiment. In an unincentivised pilot with 7 participants the average duration for one forecast submission was 49 seconds. So, 80 seconds is an adequate timeframe to reach a intuitive forecasting decision. Compare also to Pfajfar and Zakelj (2011) providing a timeframe of 80 seconds when facing participants with three macroeconomic variables and two decisions to type in.

[13] See for example Adam (2007), Pfajfar and Zakelj (2011), Assenza et al. (2013) and Odria and Rodriguez (2013) using a similar payoff function.

# 4 Results

## 4.1 Descriptive Analysis

The experiment was conducted in the summer of 2014 at the University of East Anglia in Norwich. Each of the 37 participants (14 males and 23 females) took part in $T^-$ (20) or $T^+$ (17) only. The participants were invited using the database of the Centre for Behavioural and Experimental Social Science (CBESS). The average earnings were £8.32 for an average duration of 46 minutes.

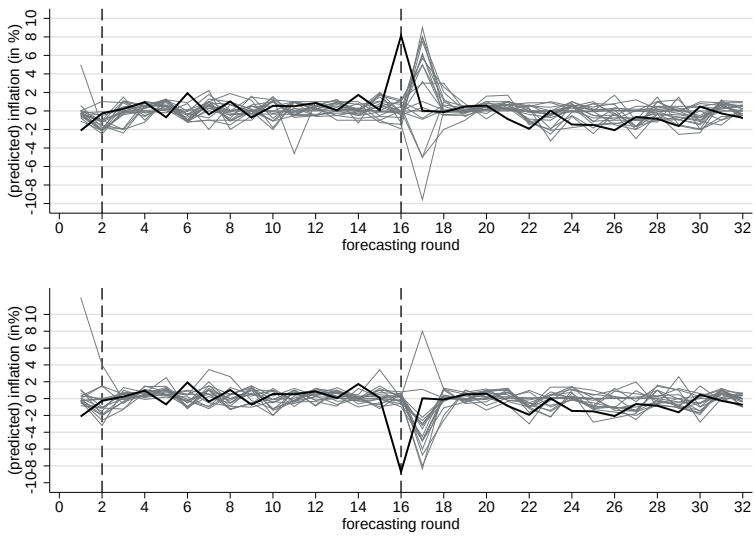
Figure 1 displays the forecasted inflation series of all participants and the real inflation trajectory. Participants' forecasts are relatively close to real inflation. In the pre-phase disagreement among participants seems to be rather low. Also, stated forecasts appear to be less volatile than the real inflation rate in both treatment groups. The spikes caused by the supply shocks are not anticipated by the participants. In the first post-shock periods forecasting disagreement increases and does not return to pre-phase level. Other than that, the shocks seem to have no lasting impact on the level of inflation forecasts. It is not clear whether the direction of shocks affects inflation forecasts differently. However, in the first post-shock period $T^-$-participants predict an acceleration in the opposite direction of the shock more frequently than in $T^+$. Besides, more people predict a mean reversion in the first post-shock period in $T^-$, whereas in $T^+$ that is the case for one participant only. So, directly after the shock it seems that the disagreement across participants in $T^-$ is higher than in $T^+$.

Table 2 summarises the average standard deviation of stated forecasts over time. Before the shock takes place the average standard deviation is 0.974 and 0.840 for $T^-$ and $T^+$, respectively. From period 17 to period 19, the standard deviation increases heavily to about 3.6. Then, the larger the distance to the shock, the lower the standard deviation. In total, standard deviation after the shock is much higher than before the shock. $T^+$-participants seem to state slightly less volatile forecasts.

Table 2: Average Standard Deviation of Stated Forecasts

| | **before** | **2 periods after** | **5 periods after** | **10 periods after** | **after** |
|---|---|---|---|---|---|
| $T^-$ | 0.974 | 3.708 | 2.421 | 1.889 | 1.616 |
| $T^+$ | 0.840 | 3.287 | 2.357 | 1.801 | 1.551 |

Figure 1: Forecasted Inflation in $T^-$ (top) and $T^+$ (bottom)




Note: Real inflation rate in black and individual forecast series in grey.

## 4.2 Individual Forecasting Taxonomy

The data collected from this experiment can be described as balanced panel data. So, for each of the 37 participants there are 30 information points over time. Due to the time series dimension of the data, *serial correlation* might be an issue. Using the test proposed by Durbin and Watson (1950) the null hypothesis of no serial correlation is rejected for almost every individual inflation forecast series at the 5%-significance level.[14] Accordingly, following Newey and West (1986) estimated standard errors are corrected, hence, statistical inference is valid.

Firstly, model (1) is fitted to the data to test for unbiasedness (U), efficiency (E) and hence rationality of stated forecasts. This is tested by means of a Wald-F-Test.[15]

Table 3 presents the results. Before the shock takes place, only one participant's forecasts are unbiased. 18.9% ($\frac{7}{37}$) of all participants build efficient forecasts. Using the Fisher-Exact Test (Chochran, 1952) differences between $T^-$ and $T^+$ are not significant (unbiasedness: $FE$, p-value= 0.541; efficiency: $FE$, p-value= 0.404). That is not surprising since the information provided is identical in both treatments before the shock.

[14]For two participants the null cannot be rejected. However, due to the nature of the task Newey-West standard errors are used also.

[15]Note, lags of inflation are not modeled in the respective OLS-regression because of the linear deterministic relationship between forecast errors and inflation, and hence perfect multicollinearity.

Table 3: Rationality Test

| | *before* | | | *after* | | |
|---|---|---|---|---|---|---|
| | **U** | **E** | **R** | **U** | **E** | **R** |
| $T^-$ | 1 | 3 | 1 | 2 | 20 | 2 |
| $T^+$ | 0 | 4 | 0 | 1 | 0 | 0 |

As a result, only one participant is found to state rational forecasts when pre-forecasting. Using the test by McNemar (1969) for differences of binary variable outcomes between paired replicates, the fraction of participants forming unbiased forecasts does not change between pre- and post-forecasting ($\chi^2_{MN} = 2.00$, p-value= 0.157). After the shock, differences between treatments are insignificant when using the Fisher-Exact Test ($FE$, p-value= 0.562). Interestingly, after the shock takes place forecasts of all $T^-$-participants can be described as efficient. This is a significant change compared to the fraction of participants forming efficient forecasts before the shock in $T^-$ ($\chi^2_{MN} = 17.00$, p-value= 0.000). But, this pattern cannot be observed in $T^+$. So, after the negative supply shock the efficiency of inflation forecasts increase heavily, whereas after a positive supply shock low-level efficiency drops down to zero. The number of participants building rational forecasts does not significantly change due to the shock and between treatment groups (pre/post-comparison: $\chi^2_{MN} = 1.00$, p-value= 1.000; $T^-/T^+$-comparison: $FE$, p-value= 0.896).

Elicited inflation expectations are barely rational, hence, alternative bounded rational models may explain intuitive forecasting. To classify which of the presented models, (2)-(9), explain individual forecasting best, estimation results are compared in line with their goodness of fit. For this purpose the *Akaike Information Criterion* (AIC) and the *Baysian Information Criterion* (BIC) are used.[16] Using Maximum-Likelihood estimation, the distance between the real and the estimated model is numerically expressed. Thus, the lower the AIC and the BIC are, the closer an estimated forecasting rule is to the real forecasting rule used by a participant (Anderson and Burnham, 2004). If the AIC-values are very close for several rules or disagree with the values of BIC, the model selection is according to the lowest BIC. The BIC penalises a specific forecasting rule according to the number of regressors incorporated (Kuha, 2004).

To differentiate between univariate AR-form forecasts (eq. (2)) and SE-forecasts (eq. (5)) a Wald-F-Test is used. If an individual forecast trajectory is classified as (2)-type forecast according to the information criteria, whether the constant and the coefficient of the AR-model are 0 and 1 are tested. If so, forecasts are considered to be SE. Among

[16]For a comparison of both criteria see Kuha (2004).

Table 4: Forcasting Taxonomy

| | *before* | | | *after* | | |
|---|---|---|---|---|---|---|
| | $T^-$ | $T^+$ | **all (in %)** | $T^-$ | $T^+$ | **all (in %)** |
| **SE** | 5 (0) | 1 (1) | 16.2 | 0 (0) | 1 (0) | 2.7 |
| **AR(1)** | 3 (0) | 3 (0) | 16.2 | 7 (2) | 7 (0) | 37.8 |
| **AR(3)** | 3 (3) | 4 (2) | 18.9 | 1 (1) | 0 (0) | 2.7 |
| **TE** | 0 (0) | 1 (0) | 2.7 | 3 (1) | 2 (0) | 13.5 |
| **AL** | 1 (0) | 2 (1) | 8.1 | 2 (0) | 2 (0) | 10.8 |
| **NE** | 7 (1) | 6 (1) | 35.1 | 4 (0) | 2 (1) | 16.2 |
| **AA** | 1 (0) | 0 (0) | 2.7 | 2 (0) | 3 (0) | 13.5 |

AR-form forecasting rules AR(1)- and AR-(3) processes are found to describe forecasting best, so, are considered in the classification procedure. Similarly, the NE-form forecasting rule is modelled such that in each case three lags of output gap and inflation are used.

Table 4 presents the taxonomy resulting from the individual forecast classification procedure. Before the shock takes place, 2.7% of the participants form either TE- or AA-form forecasts, 8.1% AL- and 16.2% SE-form forecasts. Forecast series of 16.2% and 18.9% of the participants can be described by a AR(1)- or AR(3)-process, respectively. 35.1% use a NE-rule. The population differences between both treatment groups are not significant using the test introduced by Kolmogorov (1941) and Smirnov (1948) for equality of distribution functions ($KS = 0.165$, p-value= 0.936). After the shock takes place, 40.5% of participants can be best described by using a univariate one-lag rule (either AR(1) or SE). Still a relatively high fraction are found to form NE (16.2%) and 13.5% use TE- and AA-rules, respectively. 10.8% can be described by AL and 2.7% by an AR(3)-process. We find that 63% of the participants in $T^-$ and 75% in $T^+$ change their forecasting rule before and after the shock. So, forecasting rule switching is evident at the 1%-significance level ($\chi^2_{MN} = 25.00$, p-value= 0.000). Differences between treatments are not significant ($FE$, p-value= 0.481). The test on marginal homogeneity of population replicates (Stuart,1955; Maxwell, 1970) reports a significant difference between pre- and post- forecasting taxonomies at the 10%-significance level ($\chi^2_{SM} = 11.45$, p-value= 0.075). However, when testing differences between pre- and post-forecasting taxonomies for both treatments separately results are mixed. In $T^-$ taxonomies before and after the shock are not significantly different ($\chi^2_{SM} = 5.00$, p-value=0.544). In $T^+$ the shock affects the forecasting taxonomy ($\chi^2_{SM} = 11.10$, p-value= 0.050). Thus, the qualitative picture how participants form their forecasts changes due to a positive supply shock rather than a negative supply shock. The use of simple univariate one-lag rules increases significantly

in the $T^+$-group due to the shock ($\chi^2_{MN} = 3.00$, p-value= 0.089). This is not the case for $T^-$ ($\chi^2_{MN} = 0.14$, p-value= 1.000). So, after the positive supply shock participants change to an even simpler forecasting rule. Other sytsematic switching patterns cannot be observed.

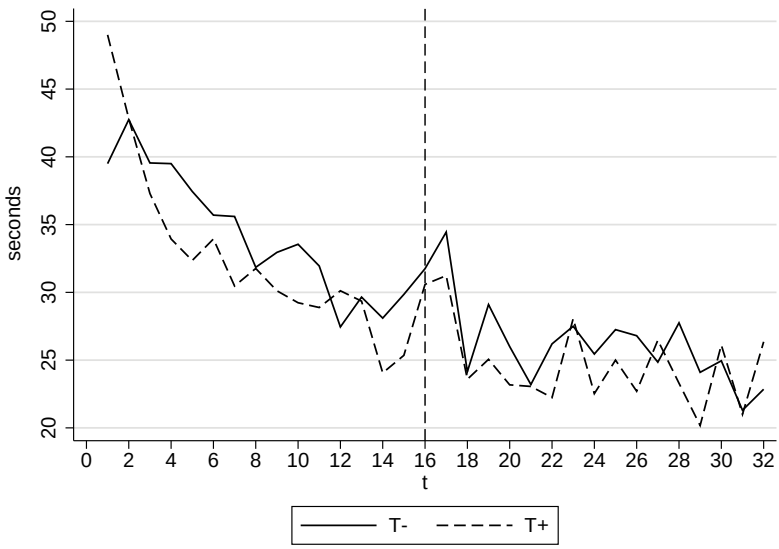
### 4.3 Dynamic Analysis

As emphasised by Pfajfar and Zakelj (2014), participants switch among several forecasting rules even though there is no structural break inducing them to do so. In table 4, the numbers in parentheses indicate how many participants' forecast series do not pass a parameter stability test. For each participant the best fitting model is stressed by a Cumulative-Sum Test (CUSUM) checking whether the coefficient of the respective regressors vary over time. For instance, before the shock takes place 5 out of 7 used AR(3)-rules are found to be unstable over time at a significance level of 5%. Such instability could give an hint of forecasting rule switching behaviour during pre- and post-forecasting. For every participant's forecasting series which does not pass the CUSUM-test, all models are reestimated in a rolling window regression. Here, the rules are estimated using a data window of 5 periods. In the first iteration step, forecasts from period 3 to period 7 are fitted to models (2)-(9) and classified according to their AIC and BIC. In the next step, the estimation window reaches from period 4 to period 8 and the classification procedure is done a second time. This continues on until the last period is reached.[17] If two consecutive window estimations are classified by the same rule, the participant is considered to use the respective rule. On average, the 9 participants before and the 4 participants after the shock use 2.4 different rules, hence, switch their forecasting rule every 4th period. The Sign-Rank Test as described by Wilcoxon (1949) reports no significant difference between the number of rules used before and the respective number after the shock ($Z = 1.080$, p-value= 0.2803). Also, there is no significant difference between treatments ($FE$, p-value= 0.772).

The analysis of differences between pre- and post-forecasting relies on a within-subject comparison. Thus, treatment effects are possibly due to ordinary learning behaviour rather than being due to the shock experience. Frequently the time needed by a participant to submit a decision is used to make statements about learning behaviour (see Sitzia and Zizzo, 2011). Figure 2 shows the average time needed to submit the inflation forecast for both treatments. It can be seen that there are no large differences between treatments. Both trajectories can be interpreted as downward-sloping curves. The more decisions have been made the less time is needed; potential evidence for ordinary learning. The average time needed spikes in the first period after the shock. However, in period

[17] A margin of four data points at the end of the time series are not considered in the estimation because of the window size of 5 periods.

Figure 2: Average Time Needed to Submit a Forecast




18 the pre- forecasting pattern continues. It seems that both effects could be present, ordinary learning and shock effects. As an attempt to distinguish between both effects, all models are fitted to the data again but without differentiating between before and after shock data. The rules are estimated using a rolling window procedure (5 periods), so, changes in the coefficients over time can be displayed and a qualitative analysis of adjustment processes is possible.[18] As a main result, there is evidence that few of the rule changes in the pre/post-forecasting comparison could be due to ordinary learning; however, there are also indications for systematic variations in coefficient adjustment after the shocks.

## 4.4 Aggregated Analysis and Other Robustness Checks

Menzies and Zizzo (2006) state that "[...] rational expectations allows for mistakes as long as they are not systematic mistakes" (p. 3). Whether deviations from rationality in the experiment here are systematic can be tested by aggregating the elicited individual-level forecasts. Following Dangerfield and Morris' (1992) Bottom-Up Forecasting Aggregation, an average of all individual forecasts is formed. Similar to the individual-level analysis unbiasedness, efficiency and rationality are tested. Then the forecasting rule that fits the

[18]For a detailed analysis see appendix.

data best is estimated, again using AIC and BIC.

The results are presented in table 5. Still, forecasts are not rational. So, deviations from rationality appear to be systematic. Using the average forecasting of the pooled sample (both $T^-$ and $T^+$), forecasts become efficient after the shock. However, this is confirmed in the negative supply shock treatment only. The anchoring-and-adjustment heuristic fits the data of all participants and for all periods best. As is the case when analysing $T^-$ only. We see that all before and after comparisons come along with a change in the rule fitting the forecast trajectories best. Whereas the AA describes only a few individual forecast series, aggregated data fits well.

From individual forecasting analysis we found that efficiency in forecasting after a negative supply shock increases heavily and forecasting rules describing participants inflation forecasts best change due to a shock. Besides, after the shock $T^+$-participants use simpler rules (i.e. rules with less regressors) more frequently. As it is common practice in experimental economics the robustness of these treatment effects are checked by regression analysis.[19] By doing so, all three findings can be indicated by a dummy, such that if a participant uses the simplest rule (AR(1) or SE) the dummy $simple_rule$ is 1 otherwise it is 0. Similarly, if a participant forms efficient forecasts the dummy $efficient = 1$. $different = 0$ can be interpreted as indicating the pre-forecasting rule of a participant.

Table 5: Aggregated Forecasting

| | test | **before** | **after** | **total** |
|---|---|---|---|---|
| all | $U$ | no | no | no |
| | $E$ | no | yes | yes |
| | $R$ | no | no | no |
| | *rule* | AR(3) | AA | AA |
| $T^-$ | $U$ | no | no | no |
| | $E$ | no | yes | yes |
| | $R$ | no | no | no |
| | *rule* | SE | AA | AA |
| $T^+$ | $U$ | no | no | no |
| | $E$ | no | no | no |
| | $R$ | no | no | no |
| | *rule* | AA | AL | SE |

[19]Note, regression analysis relies on assumptions which may not be fulfilled by the collected data at hand (e.g. normality). Partly this is due to the relatively small sample size. As a result treatment effects are less likely to be confirmed. That is why results of the following robustness checks are interpreted as an indication of either weak or rather strong treatment effects.

Table 6: Random Effects Probit Models

| | simple_rule | efficient | different |
|---|---|---|---|
| shock | 0.600 | | 3.962** |
| | (0.469) | | (1.903) |
| $T^-$ | -0.0539 | -1.255*** | -0.191 |
| | (0.484) | (0.440) | (0.559) |
| shock*$T^-$ | -0.811 | 2.839*** | |
| | (0.649) | (0.641) | |
| gender | -0.714* | -0.280 | -1.256 |
| | (0.390) | (0.415) | (0.795) |
| eco_student | -0.215 | 0.783 | -0.850 |
| | (0.637) | (0.704) | (0.988) |
| literacy1 | 0.382 | 0.103 | -1.275 |
| | (0.521) | (0.525) | (0.990) |
| literacy2 | 1.272* | -0.736 | 0.074 |
| | (0.673) | (0.809) | (1.050) |
| charts | -0.163 | -0.0926 | -0.163 |
| | (0.333) | (0.355) | (0.537) |
| income_crisis | -0.126 | 0.0916 | 0.075 |
| | (0.429) | (0.445) | (0.650) |
| age | -0.0239 | 0.0113 | -0.066* |
| | (0.0151) | (0.0122) | (0.036) |
| Subjects | 37 | 37 | 36 |

Newey-West standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

If the respective participant changes his/her model when post-forecasting $different = 1$, otherwise $different = 0$. Since these dummies are the dependent variable in the respective regression a Probit Model approach can be used. When investigating *simple_rule*, $efficient$ and $different$, a before/after-comparison is made, so the panel structure of the data is used. Hence, a Random Effects Probit Model is appropriate.[20]

The treatment effects are stressed against socio-demographic and task-related characteristics of the participants. Earlier studies report many factors influencing heterogeneity in the level of stated inflation forecasts. For example, Souleles (2004) finds that females tend to have higher inflation expectations than men. Ehrmann and Pfajfar (2014) report that people exhibiting pessimistic attitudes state less accurate inflation forecasts. Burke

[20]The most appropriate model to examine the probability of using a specific rule before and after the shock would be a Mixed Effects Logit Model with Random Effects. As Pforr (2011) points out such a model has not yet been fully developed.

and Manz (2011) show that race, education or gender differences in forecast are no longer significant when controlling for *economic literacy*, i.e. the degree of knowledge about economic issues. Potentially, these factors influence forecasting rule selection, efficiency in forecasting and shock responses as well. Subsequently, information was collected in a post-experimental questionnaire. The independent variable $T^-$ indicates whether a participant attended to treatment $T^-$ ($T^- = 1$) or in $T^+$ ($T^- = 0$). $shock = 1$ indicates after the shock and $shock = 0$ before the shock. The interaction variable $shock^*T^-$ displays differences between treatments after the shock.[21]

Table 6 presents the regression results. We see that the probability of a participant forming efficient forecasts is significantly lower in $T^-$ compared to $T^+$ after the shock. However, the suggestion that $T^+$-participants switch to a simpler rule after a positive supply shock takes place is not robust. Shocks do not influence the probability of using a simple forecasting rule. Instead, gender and literacy are significant determinants. Females and more literate participants use AR(1) or SE forecasting rules with a higher probability. When stressing that participants change their forecasting rule due to the shock, treatment effects are robust. At a significance level of 5% the shock induces participants to deviate from their pre-phase forcasting rule when post-forecasting. Differences between treatments are not significant.

# 5 Discussion

In the absence of a shock, findings from this experiment confirm previous experimental results. According to unbiasedness and efficiency requirements participants barely form rational forecasts. And, individual forecasting taxonomies are comparable to those found in earlier studies (Odria and Rodriguez, 2013; Roos and Luhan, 2013). As suggested by Fuster et al. (2010) most participants (35%) use bivariate analysis to form forecasts. The emphasis put on the AA heuristic is not confirmed. However, if individual forecasts are aggregated, AA appears to describe forecasts better. Furthermore, about 19% of participants are found to switch their forecasting rule every 4th period which is similar to findings presented by Pfajfar and Zakelj (2014, p. 148).

The answer to the question how macroeconomic shocks affect intuitive inflation forecasting is ambigiuos. That is because forecasting behaviour reacts differently according to the direction of the shock. After both shocks, participants change their forecasting rule. Though after a negative supply shock, efficiency of forecasts increases strongly, whereas efficiency drops down to zero after a positive supply shock. Qualitatively, the forecasting taxonomy does not change after a negative but does change after a positive supply shock.

[21] Detailed variable declaration of the other regressors can be found in appendix C.

Here, participants are found to use forecasting rules incorporating less regressors with a higher probability. However, this effect is rather weak since it does not pass a robustness check.

It can be argued that results from the negative supply shock treatment provide evidence for a *salience heuristic*. As Gnan et al. (2010) describe "[...] people only pay attention to information that stands out" (p. 31). When experiencing the shock, it becomes more salient to the participants that inflation and output gap are linked and hence they use information about output gap to avoid larger forecast errors. More forecast series are deemed to be efficient. An alternative explanation is that participants increase effort put into forecasting due to the shock and study output gap for valuable information more intensively. Similarly, after a negative demand shock Odria and Rodriguez (2013) report "[...] that subjects are nonlinearly-inattentive" (p. 65), i.e. information from output gap is disregarded if the variable behaves 'normal', however, any conspicuos observations trigger participants to rely more on output gap.

Findings from the positive shock scenario contradict this interpretation. Participants seem to place less emphasis on output gap to form inflation forecasts. This is evidence for an asymmetry of shock responses in intuitive forecasting; a negativity bias. The bias could be due to stronger loss aversion in the negative shock scenario amplifying effort and saliency effects (Soroka, 2006).[22] Also, the positive supply shock might be less familiar to participants. The usage of simpler rules could suggest that participants are unsure whether output gap is useful to predict future inflation. Actually, two participants in the $T^+$-treatment answered the question "How did you predict inflation?" by saying that they were surprised about periods of strong deflation. Overall, the shock could increase saliency of the relation between output gap and inflation but the direction of the shock may induces adverse effects.

We have to keep in mind that results from this experiment are very specific. The forecasting environment presented to the participants is artificially constructed, so, time series properties (e.g. variance) are characteristics of this specific model calibration. In terms of Bernasconi et al. (2009), "[...] within this framework one can study expectations only for the given abstract model" (p. 254). For example, the rather weak shock effects on the time series properties of forecasts might be due to the fact that the shocks were short in their duration (one period) and directly midway through the session. This could have lead to participants anticipating that the forecasting task was about a before/after-shock comparison and hence assumed inflation to continue as observed before the shock.[23]

[22] Strictly speaking, participants loose no money, but earn less. So, induced loss aversion can be interpreted as relative to participants' past earnings.

[23] Indeed, there were trade-offs between the number of trial periods, the position of the shock and the amount of data points over time, hence, statistical power.

In terms of external validity, the use of non-real-world economic data or isolation of individuals are features decreasing comparability to the world outside the lab. Also, in the experiment the shock effects have a weaker impact on individuals because participants couldn't loose any money as it is the case in reality. Furthermore, internal validity could have been improved by expanding the trial phases to rule out learning effects.

This experiment opens several avenues for future research. Firstly, varying features to create a mosaic of factors influencing individual forecasting could help to understand why people change their forecasting behaviour due to a shock. Next to varying model characteristics one could observe the effects of market feedback or group interaction in one experimental environment. Other than that, it is worth analysing whether the kind of shock (demand or supply shock) alter the relationships subjects assume when building NE. Besides, shock effects on subjective uncertainty feelings, risk attitudes, pessimistic moods, fatalism and their impact on intuitive forecasting have to be investigated. Does a shock-induced change of risk attitudes comes with a switch of forecasting rules? Secondly, it has not yet been analysed whether socio-demographic characteristics of participants are prepositions for which forecasting rule a subject uses. Finally, shaping effects on forecasting could be incorporated in agent-based or other macroeconomic models with endogenous feedback mechanisms. As an example, DeGrauwe (2011) models economic agents in a Dynamic Stochastic General Equilibrium (DSGE) model using simple biased rules to forecast future output and inflation. What if agents are modelled as they endogenously change their forecasting behaviour due to eye-catching observations in past periods?

# 6 Conclusion

It has become a fact in economics that macroeconomic shocks influence individual-level behaviour. This study demonstrated that to some extent this is true for intuitive inflation forecasting, as well. As a main contribution to the literature we found that shocks affect whether agents form efficient forecasts and induce them to switch their forecasting rules. These findings, however, vary according to the direction of the shock. The statistical analysis suggested that a negative supply shock induced participants to form more efficient forecasts, whereas a positive supply shock didn't. We interpreted this finding as evidence of a negativity bias. The negative shock scenario was more familiar to participants and could have created stronger loss aversion. So, participants put more effort into forecasting and saliency of the link between output gap and inflation increase. After the positive supply shock, participants used simpler rules to those used before the shock. Using regression analysis, however, this effect is found to be not robust. Results are in line with earlier studies investigating the effects of macroeconomic news shocks on inflation

perception reported in surveys. Overall, forecasting before the shock fitted best with a NE and after the shock with AR(1)-form forecasts. When aggregating the data, deviations from the rationality paradigm were robust and AA was found to describe forecasting much better than it did for individual forecasting. Effects on time series properties of forecasts were rather short-term. Future research should explore the question of how intuitive inflation forecasts are endogenously influenced in a macroeconomic context further.

## Acknowledgement

This paper is based on my master thesis yielding to the MSc in Experimental Economics at the University of East Anglia, Norwich (UK). First, I would like to thank my supervisor Enrique Fatas for his valuable comments and the enlightening discussions. Also I would like to thank Ailko van der Veen for his excellent organisational help, and all seminar participants at the conference "New Economic Thinking on an Evolving Financial System 2014" in Amsterdam. I express my gratitude to Wolfgang Luhan introducing me to the arts of journal submission. The financial funding of the School of Economics at UEA is greatfully acknowledged. Finally thanks to Daniel Zizzo for providing many inspirations during my studies.

# References


Adam, K. (2007), "Experimental evidence on the persistence of output and inflation", *Economic Journal*, vol. 117, pp. 603-635.

Anderson, D.R. and K.P. Burnham (2004), "Multimodel inference: Understanding AIC and BIC in model selection", *Sociological Methods Research*, vol. 33, pp. 261-304.

Andolfatto, D., S. Hendry and K. Moran (2008), "Are inflation expectations rational?", *Journal of Monetary Economics*, vol. 55, pp. 406-422.

Anufriev, M. and C. Hommes (2012), "Evolution of market heuristics", *The Knowledge of Engineering Review*, vol. 27, pp. 255-271.

Assenza, T., P. Heemeijer, C. Hommes and D. Massaro (2013), "Individual expectations and aggregate macro behavior", Tinbergen Institute Discussion Paper No. 13-016/II.

Bernasconi, M., O. Kirchkamp and P. Paruolo (2009), "Do fiscal variables affect fiscal expectations? Experiments with real world and lab data", *Journal of Economic Behavior & Organization*, vol. 70, pp. 253-265.

Beshears, J., J. Choi, A. Fuster, D. Laibson and B.C. Madrian (2013), "What goes up must come down? Experimental evidence on intuitive forecasting", *American Economic Review*, vol. 103, pp. 570-574.

Bovi, M. (2014), "Shocks and the expectations formation process. A tale of two expectations", Working Paper, ISTAT - Italian National Institute for Statistics, Department of Forecasting and Econometric Analysis.

Branch, W. (2004), "The theory of rational heterogeneous expectations: Evidence from survey data on inflation expectations", *The Economic Journal*, vol. 127, pp. 264-295.

Brock, W.A. and C. Hommes (1997), "A rationale route to randomness", *Econometrica*, vol. 65, pp. 1059-1129.

Burke, M.A. and M. Manz (2011), "Economic literacy and inflation expectations: Evidence from a laboratory experiment", Public policy Discussion Papers No. 11-8, Federal Reserve Bank of Boston.

Cameron, L. and M. Shah (2013), "Risk-taking behavior in the wake of natural disasters", NBER Working Paper Series No. 19534, National Bureau of Economic Research.

Carroll, C.D. (2003), "Macroeconomic expectations of households and professional forecasters", *Quarterly Journal of Economics*, vol. 118, pp. 269-298.

Cochran, V.J. (1950), "The comparison of percentages in matched samples", *Biometrika*, vol. 37, pp. 256-266.

Coibion, O. and Y. Gorodnichenko (2012), "What can survey forecasts tell us about information rigidities?", *Journal of Political Economy*, vol. 120, pp. 116-159.

Dangerfield, B.J. and J.S. Morris (1992), "Top-down or bottom-up: Aggregate versus disaggregate extrapolations", *International Journal of Forecasting*, vol. 8, pp. 233-241.

DeGrauwe, P. (2011), "Animal spirits and monetary policy", *Economic Theory*, vol. 47, pp. 423-457.

Durbin, J. and G.S. Watson (1950), "Testing for serial correlation in least squares regression", *Biometrika*, vol. 37, pp. 409428.

Ehrmann, M. and D. Pfajfar (2014), "Consumer attitudes and the epidemiology of inflation expectations", Discussion Paper No. 2014-029, Tilburg University.

Evans, G.W. and S. Honkapohja (2001), "Learning and expectations in macroeconomics", Princeton University Press.

Fischbacher, U. (2007), "z-Tree: Zurich toolbox for ready-made economic experiments", *Experimental Economics*, vol. 10, pp. 171-178.

Fuster, A., D. Laibson and B. Mendel (2010), "Natural expectations and macroeconomic fluctuations", *Journal of Economic Perspectives*, vol. 24, pp. 67-84.

Garz, M. (2013), "Unemployment expectations, excessive pessimism, and news coverage", *Journal of Economic Psychology*, vol. 34, pp. 156-168.

Giuliano, P. and A. Spilimbergo (2014), "Growing up in a recession", *Review of Economic Studies*, vol. 81, pp. 787-817.

Gnan, E., J. Langthaler and M.T. Valderrama (2011), "Heterogeneity in euro area consumers' inflation expectations: Some stylized facts and implications", *Monetary Policy & the Economy*, vol. 11, pp. 43-66.

Guiso, L., P. Sapienza and L. Zingales (2013), "Time varying risk aversion", Working Paper 19284, National Bureau of Economic Research.

Hey, J.D. (1994), "Expectations formation: Rational or adaptive or...?", *Journal of Economic Behavior & Organization*, vol. 25, pp. 329-349.

Holbrook, A.L., J.A. Krosnick, P.S. Visser, W.L. Gardner and J.T. Cacioppo (2001), "Attituted toward presidential candidates and political parties: Initial optimism, inertia first impressions, and a focus on flaws", *American Journal of Political Science* , vol. 45, pp. 930-950.

Hommes, C. (2011), "The heterogeneous expectation hypothesis: Some evidence from the lab", *Journal of Economic Dynamics & Control*, vol. 35, pp. 1-24.

Kahneman, D. and A. Tversky (1974), "Judgment under uncertainty: Heuristics and biases", *Science*, vol. 185, pp. 1124-1130.

Kolmogorov, A. (1941), "Confidence limits for unkown distribution function", *Annals of Mathematical Statistics*, vol. 12, pp. 461-463.

Kuha, J. (2004), "AIC and BIC - Comparisons of assumptions and performance", *Sociological Methods Research*, vol. 33, pp. 188-229.

Lamla, M.J. and S.M. Lein (2014), "The role of media for consumers' inflation expectation formation", *Journal of Economic Behavior & Organization*, vol. 106, pp. 6277.

Lovell, M.C. (1986), "Tests of the rational expectations hypothesis", *The American Economic Review*, vol. 76, pp. 110-124.

Lucas, R.E. (1972), "Expectations and the neutrality of money", *Journal of Economic*

*Theory*, vol. 4, pp. 103-124.

Malmendier, U. and S. Nagel (2011), "Depression babies: Do macroeconomic experiences affect risk-taking?", *Quarterly Journal of Economics*, vol. 126, pp. 373-416.

Malmendier, U. and S. Nagel (2013), "Learning from inflation experiences", Stanford University and UC Berkeley Working Paper.

Mankiw, N., R. Reis and J. Wolfers (2004), "Disagreement about inflation expectations", *NBER Macroeconomics Annual 2003*, vol. 18, pp. 209-248.

Maxwell, A.E. (1970), "Comparing the classification of subjects by two independent judges", *The British Journal of Psychatry*, vol. 116, pp. 651-655.

McNemar, Q. (1969), "Psychological statistics", 4th edition, J. Wiley.

Menzies, G.D. and D.J. Zizzo (2006), "Rational expectations", Working Paper.

Muth, J.F. (1961), "Rational expectations and the theory of the business cycle", *Journal of Monetary Economics*, vol. 29, pp. 315-335.

Necker, S. and M. Ziegelmeyer (2014), "Household risk taking after the financial crisis", Working Paper No. 85, Banque Centrale Du Luxembourg.

Newey, W.K. and K.D. West (1987), "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica*, vol. 55, pp. 703-708.

O'Connor, C. (2012), "Using social representations theory to examine lay explanation of contemporary social crises", *The Journal of Community & Applied Social Psychology*, vol. 22, pp. 453-469.

Odria, L.R.M. and G. Rodriguez (2013), "Inflation expectations formation in the presence of policy shifts and structural breaks: An experimental analysis", *The Journal of Socio-Economics*, vol. 44, pp. 59-67.

Pfajfar, D. and E. Santoro (2010), "Heterogeneity, learning and inflation stickiness in inflation expectations", *Journal of Economic Behavior & Organization*, vol. 75, pp. 426-

444.

Pfajfar, D. and B. Zakelj (2011), "Inflation expectations and monetary policy design: evidence from the laboratory", Discussion Paper 2011-091, Tilburg University, Center for Economic Research.

Pfajfar, D. and B. Zakelj (2014), "Experimental evidence on inflation expectation formation", *Journal of Economic Dynamics & Control*, vol. 44, pp. 147-168.

Pforr, K. (2011) "Implementation of a multinomial logit model with fixed effects", German Stata Users' Group Meetings 2011.
Available at: http://econpapers.repec.org/paper/bocdsug11/03.htm.

Roos, M.W.M. and W.J. Luhan (2013), "Information, learning and expectations in an experimental model economy", *Economica*, vol. 80, pp. 513-531.

Roos, M.W.M. and U. Schmidt (2011), "The importance of time-series extrapolation for macroeconomic expectations", *German Economic Review*, vol. 13, pp. 196-210.

Sitzia, S. and D.J. Zizzo (2011), "Does product complexity matter for competition in experimental retail markets?", *Theory and Decision*, vol. 70, pp. 65-82.

Smirnov, N.V. (1948), "Table for estimating the goodness of fit of empircial distributions", *Annals of Mathematical Statistics*, vol. 19, pp. 279-281.

Sorensen, P.B. and H.J. Whitta-Jacobsen (2010), "Introducing advanced macroeconomics", 2nd edition, Mc-Graw-Hill Higher Education.

Soroka, S.N. (2006), "Good news and bad news: Asymmetric responses to economic information", *Journal of Politics*, vol. 68, pp. 372-385.

Souleles, N.S. (2004), "Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the michigan consumer sentiment surveys", *Journal of Money, Credit, and Banking*, vol. 36, pp. 40-72.

Stuart, A. (1955), "A test for homogeneity of the marginal distributions in a two-way classification", *Biometrika*, vol. 42, pp. 412-416.

Trichet, J.-C. (2010, November 18), "Reflections on the nature of monetary policy non-standard measures and finance theory", Opening address at the ECB Central Banking Conference, Frankfurt.
Available at: https://www.ecb.europa.eu/press/key/date/2010/html/sp101118.en.html. Accessed 19/08/2014.

**Appendices**

**A**: Additional Figures

**B**: Experimental Material

**C**: Robustness Check Variable Description

# Appendix A

Figure 3: Forecasting Data in $T^-$ (left) and $T^+$ (right)

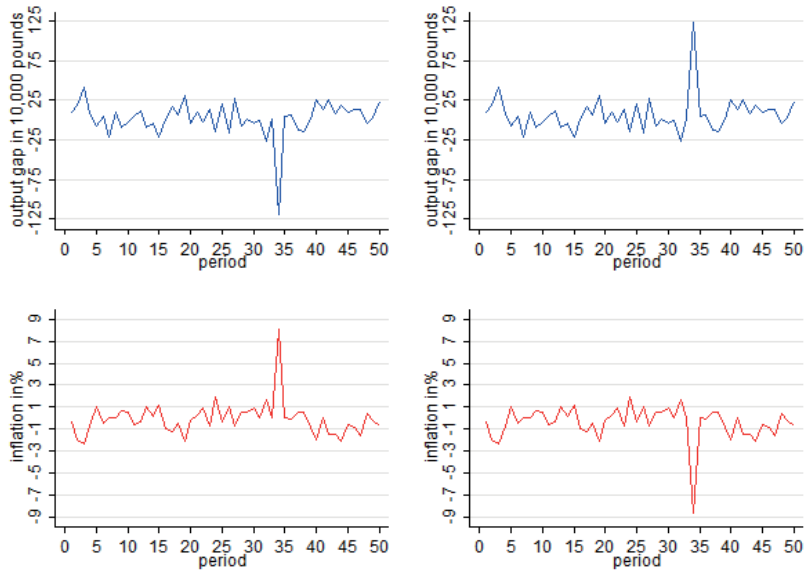



Figure 3 presents the simulated data to be forecast by participants. It can be seen that due to the shocks in period 34 inflation and output gap conspicuously spike. Before and after these spikes trajectories are comparable.

Figure 4: Example Forecasting Screen

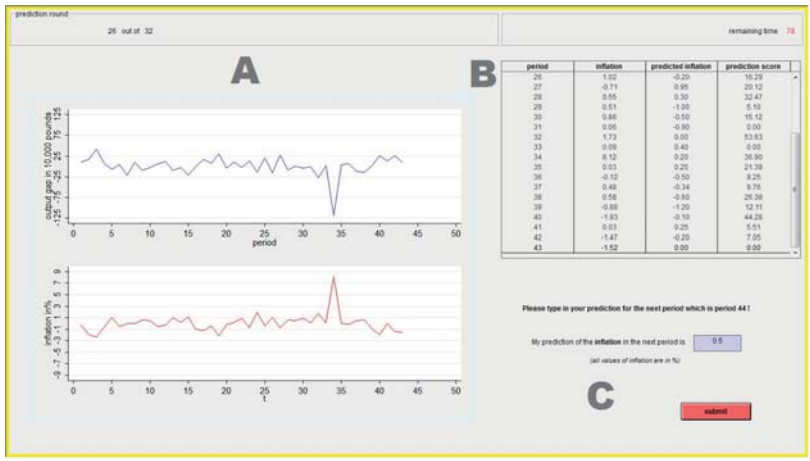
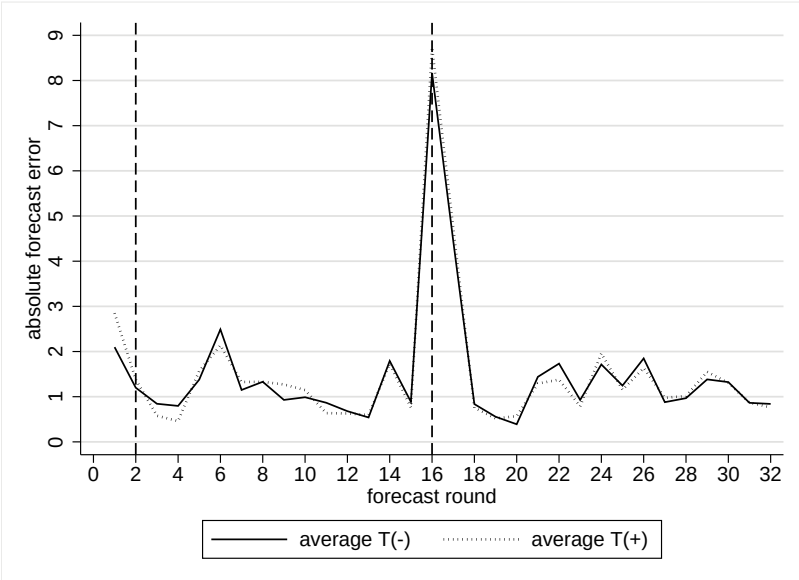


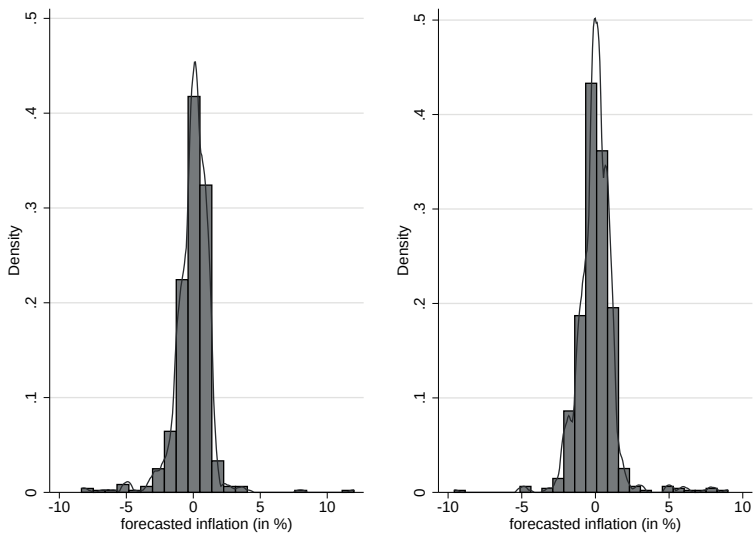

Figure 4 presents an example forecasting screen in one of the 32 periods. Charts (A) and a data list on participants' desks provide subjects with 18 past values of inflation and output gap. The next value to forecast is on the right hand side in the inflation chart. In the next round the new real value of inflation and output gap are updated. On the right, subjects see their prediction history (B). The real values of inflation are presented from period to period. The forecasting interface (C) asks subjects to type in their inflation forecast for the next period. After submitting a forecast, real inflation rates, forecast mistakes and the respective prediction scores are presented on a performance feedback screen and then added to the prediction history.

Figure 5: Average Absolute Forecast Error Dynamic




In figure 5, it can be seen that the accuracy of forecasts before and after the shock does not seem to change. It is worth noting that on average both treatment groups appear to forecast comparably accurately, before and after the shock. After only two post-shock periods the absolute forecast error reaches its pre-phase level. Similarly to figure 3, there seems to be no lasting adverse effect on participants' forecasting performance due to a one-period shock. Still, no differences between treatment groups are notable after the shock. Furthermore, one could have expected decreasing forecast errors over time due to learning effects. If anything, such a pattern can be observed in the transition of the trial periods to the actual forecasting periods for few participants. Generally, typical learning patterns are not observable.

Figure 6: Histogramms of Inflation Forecasts for $T^-$ (left) and $T^+$ (right)




The histograms in figure 6 report the information that participants predict the inflation to be 0% or around 0% most frequently. Firstly, this shows that participants may orientate to the average of the real inflation rate (0.0203%). Secondly, forecasting future inflation of 0% might have some focal power. Comparing histogramms of $T^-$ and $T^+$, no clear differences can be reported.

Figure 7: Rolling Estimation of Best Rule Coefficients

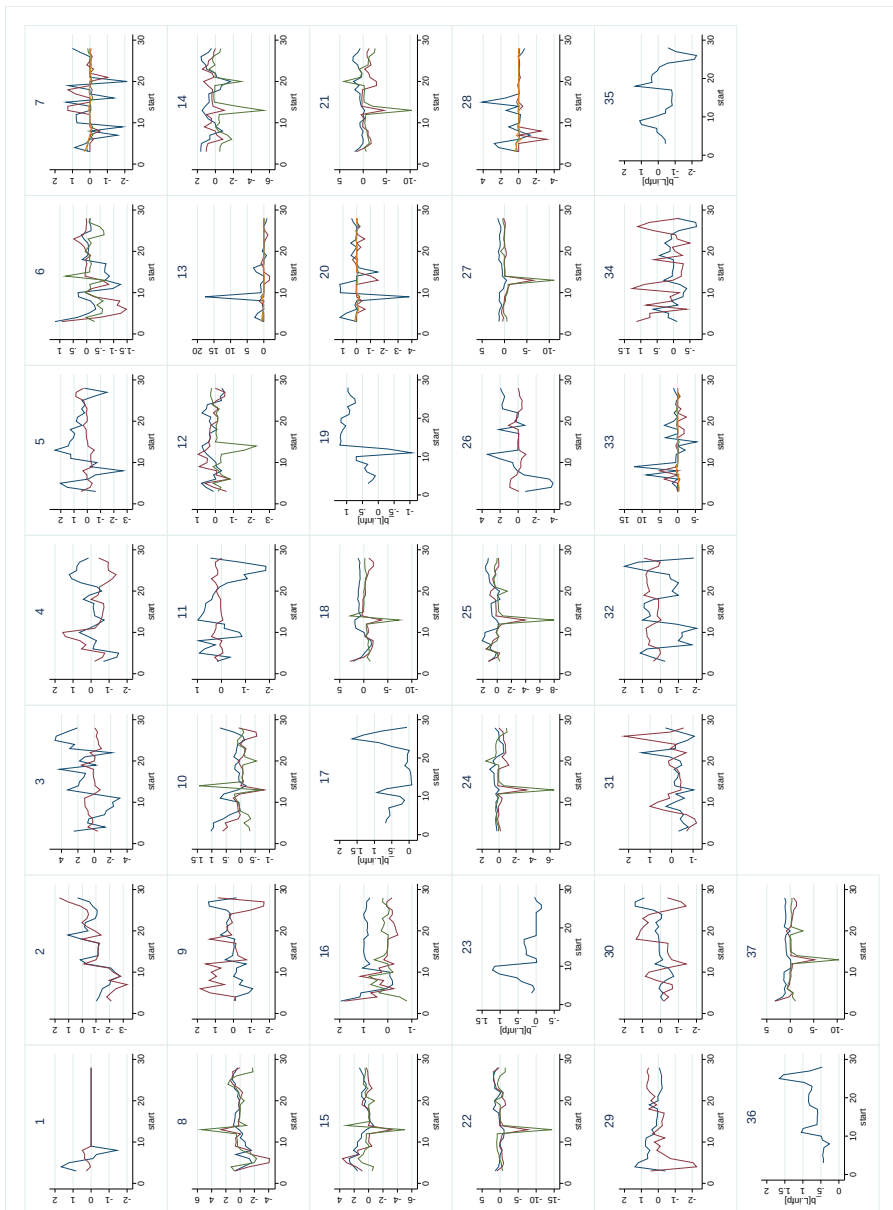


Figure 7 shows rolling estimations of the best rule coefficients for each of the 37 individuals. Note, subjects 1-20 attended to $T^-$ and subjects 21-37 to $T^+$. We can observe four different patterns: no notable changes in coefficient development, changes around the shock period and systematic changes after the shock period with either increasing or decreasing coefficient volatility. Subject number 1 cannot be classified according to one of these patterns, that is because this particular subject starts reporting 0% from period 9 on. For instance, subjects 2, 7, 23 or 29 seem to not be affected by the shock but adjust their coefficients systematically over time. Here, ordinary learning may be stronger than any shock effect or the shock is assumed to be an outlier. While period 17 is in the estimation window (periods 13-20), subjects 14, 22 or 32 display strong parameter changes. These participants return to their former pattern from period 20 on. Among others, systematic changes in the coefficient movements compared before and after the shock can be observed for subjects 3, 5, 11, 20 or 28. For example, subject 20's coefficient volatility stabilises after the shock (from period 13 on). In contrast, for subject 11 volatility is increasing.

# Appendix B

# Experimental Instructions

## Introduction

Welcome to this experiment! This is an experiment in the economics of decision making. The instructions are simple. If you follow them carefully and make good decisions, you can earn a considerable amount of money. You are awarded 2 pounds for showing up on time. Your additional earnings depend on your decisions and chance. You will be paid in cash at the end of the experiment. We will proceed as follows. First, I will read these instructions. After a short questionnaire, you can click "START" on your computer screen. If you have any questions during the experiment, please raise your hand, then someone will come to assist you. No talking during the experiment! Please remain seated until the experimenter ends the session.

## Experiment

You participate in this experiment as an agent in a fictitious economy. You work in a firm for which you will predict future values of an economic variable –namely *inflation*. Your reward per period depends on the accuracy of your predictions. The experiment consists of 32 periods. All rules and the economy you are deciding in remain the same. Neither other participant's predictions nor your predictions about the future values affect the real development of inflation. As in reality, the economy can be hit by unpredictable events, such that inflation fluctuates around.

### A. The fictitious economy

The economy you are participating in will be described by two variables: *output gap* and *inflation*. Note, you have to forecast inflation only, but you will get information about both variables.

- The output or Gross Domestic Product (GDP) of the fictitious economy displays the value of all currently produced goods and services in 10,000 pounds. As in reality the output fluctuates around a trend. Such that sometimes output is above its trend level (boom) or under its trend level (bust).

- The output gap measures the distance between the trend level of output and the current level of output. If the output gap is greater than 0, it means that the economy is producing more than at its trend. We say that the economy is in a boom. If negative, less than on the trend is produced. Here, the economy is in a bust.

- *Inflation* measures the general rise in prices for goods and services in the economy (in %). So, inflation describes price changes of all goods and services together. If inflation is greater than 0, prices in general are increasing, if negative, prices are decreasing.

- In each period output gap depends on inflation, past values of output gap and unpredictable events. And, inflation depends on output gap, past values of inflation and unpredictable events. So, both variables are connected.

## B. Your task

- You will enter the economy in period 18.
- Your task consists in predicting the value of inflation in the next period.
- For each decision you have 80 seconds.
- You will do this in 32 periods, hence from period 18 to period 50.
- Importantly, your predictions in the first two periods are trials aiming to make you familiar with the economy and the computer program. In both trial periods, predictions earn you no money.

## C. Your rewards

- Your reward after the experiment has ended depends on the accuracy of your predictions in every period.
- The accuracy is measured by the absolute prediction error. The absolute prediction error is the distance between your submitted prediction and the real value of inflation in the predicted period.
- The more accurate your prediction is, the lower the absolute prediction error.
- According to your accuracy, you get a *prediction score* in experimental points for every prediction. The table below gives you the relation between the absolute prediction error and the respective prediction score.

| **Absolute prediction error** | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 | ≥2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Prediction score** | 100 | 50 | 33/3 | 25 | 20 | 16.6 | 14.3 | 12.5 | 11.1 | 10 | 9.1 | 0 |

- You see, the more accurate your prediction is, the higher is your prediction score. If your absolute prediction error is higher than or equal to 2, your prediction score is 0. Perfect predictions yield 100 experimental points.
- An Example: Imagine in a certain period you predict an inflation of 1% in the next period. Then it turns out that the actual inflation is 0.1%. Therefore your absolute error of predicting inflation is 0.9 (1-0.1=0.9) and the respective score is 10.
- Your prediction scores in experimental points from each period are summed in the end of the experiment and determine your *total prediction score* in experimental points.
- The total prediction score is exchanged by an exchange rate of 0.012. Such that 100 experimental points are worth 1.2 pounds.

**D. Information provided**

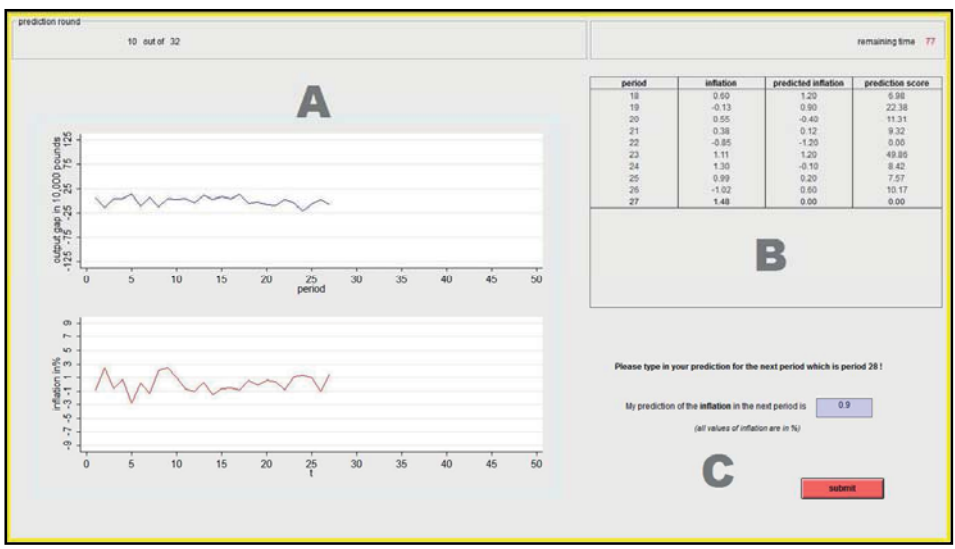
- From the first period on (period 18), you will be given past values of inflation and output gap for 18 periods back. These past values will be presented in a data list on your desk and in charts on your computer screen.

- After you submit your prediction about the next period (say period 19), you will receive a feedback. The real value of inflation in period 19, your prediction error and the respective score you made in this period will be shown to you.

**E. The computer program**

- Once you clicked the "START"-button you will be guided to the prediction screen.

- In all 32 periods you are shown the prediction screen. See here an example prediction screen as it will be presented to you:




It consists of three elements: the *charts* (A), the *prediction history* (B) and the *prediction area* (C).

- The *charts* (A) display the values of output gap (blue) and inflation (red) from period 1-50. Up to period 18 this data is the same as presented in the *past values list* on your desk. On the horizontal axes are the time periods (from 1-50); the vertical axes display the variable values. The next period to predict is always on the right hand side in the inflation chart. After predicting, the next periods' values will be updated in the charts.

- The *prediction history* (B) initially presents the value of inflation in period 18, your starting period. From round to round, your predictions about the values of inflation in the next period are added here. Such that in period 19 there will be your prediction made for period 19 in the row of period 18. That is because you made your prediction in period 18. Additionally, the prediction score achieved in each period is listed.

- The *prediction area (C)* asks you to type in the prediction of inflation for the next period. The signs "-", "." and all numbers are allowed here. Do not use %-sign. Click on submit to start calculation of your prediction scores for this period.

- After submitting your predictions, the profit screen summarizes the period. This is the respective example profit screen as it will be presented to you:

- Pressing "continue" leads you to the next prediction screen where you type in your prediction for the next period ahead.

**Questionnaires**

After reading these instructions, you will be provided with a short questionnaire to make sure that you understand everything. The instructions can remain on your desk and may be consulted during the experiment. After you raised your hand the experimenter will check your answers and provide further advice if necessary. Then you can click "START".

At the end of the experiment another questionnaire asks you how the experiment went.

Eventually, raise your hand and wait for the experimenter to pay you out; your earnings will be shown on the screen, please write this amount on the receipt on your desk.

During the experiment read the instructions on the screen carefully. These will repeat what the next step is.

## Questionnaire

Do your predictions or other participants' predictions affect the real values of output gap and inflation?

O Yes O No

Will you receive a feedback after every period?

O Yes O No

If the output gap is positive (>0) the economy produces more than its trend level.

O True O False

The inflation displays how much the price of **one single** good (say bread) increases or decreases?

O Yes O No

The more accurate your predictions are, the higher is your prediction score and hence your earnings from this experiment.

O True O False

If the deviation between your prediction and the true value of inflation is larger or equal to 2, you achieve a prediction score of 0.

O True O False

*Raise your hand when you have finished this questionnaire.*

**Data list**

| period | output gap *(in 10,000 pounds)* | inflation *(in %)* |
|---|---|---|
| 1 | 9.45 | -0.29 |
| 2 | 17.88 | -1.94 |
| 3 | 41.15 | -2.31 |
| 4 | 7.63 | -0.66 |
| 5 | -7.39 | 1.03 |
| 6 | 5.08 | -0.51 |
| 7 | -21.38 | 0.01 |
| 8 | 9.73 | 0.03 |
| 9 | -9.61 | 0.62 |
| 10 | -2.81 | 0.52 |
| 11 | 5.37 | -0.56 |
| 12 | 11.27 | -0.34 |
| 13 | -9.95 | 1.02 |
| 14 | -3.24 | 0.22 |
| 15 | -21.65 | 1.11 |
| 16 | 0.52 | -0.98 |
| 17 | 16.68 | -1.29 |
| 18 | 7.27 | -0.42 |

FOR _ _ - _

## Individual Background

Age: _____

Gender: _____

Field of Study: _____

Did you ever attended to classes on macro- or financial economics? O Yes O No

Do you know what AS-AD macro models are? O Yes O No

When making your predictions, on which information did you rely more?

O charts O numerical values (as in the data list or in the prediction history)

In which country did you spend the majority of your life?

_____

Did the economic crisis in 2008/2009 affect your personal income situation?

O Yes O No

*Answer on backside:*

Explain how you predicted inflation?

# Appendix C

**Regressors and individual background questions in the robustness check section**

| regressor | question | definition |
|---|---|---|
| *gender* | "Gender: _____" | 1 for male<br>0 for female |
| *eco_student* | "Field of Study: _____" | 1 if economics<br>0 if not economics |
| *literacy1* | "Did you ever attended to classes on macro- or financial economics? O Yes O No" | 1 if subject answered 'yes'<br>0 if subject answered 'no' |
| *literacy2* | "Do you know what AS-AD macro models are? O Yes O No" | 1 if subject answered 'yes'<br>0 if subject answered 'no' |
| *charts* | "When making your predictions, on which information did you rely more? O charts O numerical values (as in the data list or in the prediction history)" | 1 if subject answered 'charts'<br><br>0 if subject answered 'numerical values' |
| *income_crisis* | "Did the economic crisis in 2008/2009 affect your personal income situation? O Yes O No" | 1 if subject answered 'yes'<br>0 if subject answered 'no' |
| *age* | "Age: _____" | age in years |