Online Appendix for “Monetary Policy under Behavioral Expectations: Theory and Experiment”∗

Cars Hommes  Domenico Massaro  Matthias Weber

This online appendix contains material in addition to the manuscript and to the printed appendix. As the online appendix follows the printed appendix (Appendix A), it begins with Appendix B. Similarly, Figure numbers build on the numbers in the article and the printed appendix, therefore starting with Figure 13 in this online appendix.

B  Appendix: Additional Graphs from Simulations of the Macroeconomic Model

B.1  Changes in the Parameters of the Macroeconomic Model

Figure 13 shows inflation volatility as a function of the output gap reaction coefficient $\phi_y$ for the model assuming rational expectations, similarly to Figure 1a. The graph now shows multiple coefficients of $\phi_\pi$ simultaneously (from top to bottom the lines correspond to $\phi_\pi$-values of 1.4, 1.5, 1.6, and 1.7). Figure 14 shows the same graph for the behavioral model (again the lines correspond to $\phi_\pi$-values of 1.4, 1.5, 1.6, and 1.7, from top to bottom).

Figures 15-18 show inflation volatility as a function of the output gap reaction coefficient similar to Figure 1 for different parameter values of the macroeconomic model. Parameters which are not specifically mentioned in a graph are fixed to the same standard values as used for the graphs in Section 2.3. Under rational expectations, inflation volatility increases monotonously in the output gap reaction coefficient in all of the graphs, similarly to Figure 1a. Under behavioral expectations the U-shape arises in all graphs similarly to Figure 1b.

∗Citation format: Hommes, Cars, Domenico Massaro, and Matthias Weber, Online Appendix for “Monetary Policy under Behavioral Expectations: Theory and Experiment”, European Economic Review.
Figure 13: Inflation volatility as a function of $\phi_y$ for the rational model for different values of $\phi_\pi$ (from 1.4 (top line) to 1.7)

Figure 14: Inflation volatility as a function of $\phi_y$ for the behavioral model for different values of $\phi_\pi$ (from 1.4 (top line) to 1.7)
Figure 15: Inflation volatility for different values of $\phi$

Notes: This figure shows the effect of parameter $\phi_y$ on inflation volatility for different values of $\phi$ ($\phi_\pi = 1.5$ throughout).
Figure 16: Inflation volatility for different values of $\lambda$

Notes: This figure shows the effect of parameter $\phi_y$ on inflation volatility for different values of $\lambda$ ($\phi_\pi = 1.5$ throughout).
Figure 17: Inflation volatility for different values of $\rho$

Notes: This figure shows the effect of parameter $\phi_y$ on inflation volatility for different values of $\rho$ ($\phi_{\pi} = 1.5$ throughout).
Figure 18: Inflation volatility for different values of the standard deviation (of both demand and supply shocks)

Notes: This figure shows the effect of parameter $\phi_y$ on inflation volatility for different values of the standard deviation of demand and supply shocks ($\phi_x = 1.5$ throughout).
B.2 Results with Different Behavioral Models of Expectation Formation

The results are robust to variations of the parameters of the behavioral model of expectation formation we employ. Furthermore, the results are qualitatively the same for a wide variety of other behavioral mechanisms. We show two examples here. Figure 19 shows inflation, output gap, and interest volatility as a function of the output gap reaction coefficient. Expectations are not formed according to the main heuristic switching model described in Section 2.2, but according to two simpler models of behavioral expectation formation. On the left side of this figure, it is assumed that agents use a heuristic switching model similar to the one described before but including only two very simple heuristics, naive expectations which always forecast the last observation and a trend-following rule with trend-following coefficient one. On the right side, the graphs show the results from naive expectations alone (thus without any switching). Here as well, the results look similar to the ones in Figures 1 and 2.

B.3 Results with Different Starting Values of Output Gap and Inflation Forecasts

Figures 20 and 21 show graphs similar to Figure 1b for different combinations of starting values of inflation and output gap (i.e. inflation and output gap are set to these starting values in the first two periods). In all cases the U-shape arises similarly to Figure 1b.

B.4 Results with an Interest Rate Smoothing Taylor Rule

One can also modify the model to include interest rate smoothing by the central bank. Including interest rate smoothing in the Taylor rule leads to aggregate New Keynesian equations of the form below, where everything is equal to our main model, except for the monetary policy rule that includes an interest rate smoothing parameter $\mu$, with $0 < \mu < 1$. Note that if the central bank places too much weight on the past interest rate, it loses its ability to steer the economy, so that $\mu$ should not be too large (in the extreme case of $\mu = 1$, the interest rate is just a constant without any reaction to
Figure 19: Inflation, output gap, and interest rate volatility for a simple HSM of expectation formation (with switching between naive expectations and trend-following) and for naive expectations.

Notes: This figure shows the effect of parameter $\phi_y$ on inflation, output gap, and interest rate volatility for alternative models of expectation formation ($\phi_\pi = 1.5$ throughout).
Figure 20: Inflation volatility in the behavioral model for different starting values

Notes: This figure shows the effect of parameter $\phi_y$ on inflation volatility for different starting values of $y$ and $\pi$ ($\phi_\pi = 1.5$ throughout).
Figure 21: Inflation volatility in the behavioral model for different starting values

Notes: This figure shows the effect of parameter $\phi_y$ on inflation volatility for different starting values of $y$ and $\pi$ ($\phi_\pi = 1.5$ throughout).
economic activity).

\begin{align}
y_t &= y_{t+1}^e - \varphi(i_t - \bar{\pi}_{t+1}^e) + g_t \tag{B.1} \\
\pi_t &= \lambda y_t + \rho \bar{\pi}_{t+1} + u_t \tag{B.2} \\
i_t &= \text{Max}\{(1 - \mu)(\bar{\pi} + \phi_{\pi}(\pi_t - \bar{\pi}) + \phi_y(y_t - \bar{y}) + \mu i_{t-1}, 0)\} \tag{B.3}
\end{align}

Figures 22 and 23 show the volatility of inflation, output gap, and interest rate under the benchmark model of expectation formation assuming an interest smoothing monetary policy rule (all values except for the new parameter $\mu$ equal the values of the benchmark model as in the right panels of Figures 1 and 2).
Figure 22: Inflation, output gap, and interest rate volatility with interest rate smoothing

Notes: This figure shows the effect of parameter $\phi_y$ on inflation, output gap, and interest rate volatility with interest rate smoothing for different values of $\mu$ ($\phi_\pi = 1.5$ throughout).
Figure 23: Inflation, output gap, and interest rate volatility with interest rate smoothing

Notes: This figure shows the effect of parameter $\phi_y$ on inflation, output gap, and interest rate volatility with interest rate smoothing for different values of $\mu$ ($\phi_x = 1.5$ throughout).
C Appendix: Discussion of the Measurement of Volatility

In general, different simple measures of price instability, i.e. of volatility, dispersion, or distance from the target are possible. We discuss mainly two of them here. The first one is the measure that we use, $v(\pi) = \frac{1}{T} \sum_{t=2}^{T} (\pi_t - \pi_{t-1})^2$ (equivalently, one could of course take $v_1(\pi) = \frac{1}{T-1} \sum_{t=2}^{T} (\pi_t - \pi_{t-1})^2$ or even $v_2(\pi) = \sum_{t=2}^{T} (\pi_t - \pi_{t-1})^2$ if the number of periods is fixed). The second one is the mean squared deviation from the target, $msd(\pi) = \frac{1}{T} \sum_{t=2}^{T} (\pi_t - \bar{\pi})^2$. Other alternatives that one could use are the absolute deviation, $ad(\pi) = \frac{1}{T} \sum_{t=2}^{T} |\pi_t - \pi_{t-1}|$, and the standard deviation, $sd(\pi) = \frac{1}{T} \sum_{t=2}^{T} (\pi_t - \pi^{av})^2$, where $\pi^{av}$ is the average of inflation in a group taken over the whole time period. We do not discuss these measures here in detail; in general, $ad(\cdot)$ shares many features with $v(\cdot)$, and $sd(\cdot)$ shares many features with $msd(\cdot)$.

The measures $v(\cdot)$ and $msd(\cdot)$, are different in the following ways. The mean squared deviation from the target exclusively takes into account the distance to the target, not whether or not this distance is positive or negative. Figure 24 illustrates this with two example time series.

![Figure 24: Example time series](image)

The solid red line and the dashed blue line have exactly the same distance from the target in each period. However, it seems clear that the red line is much more volatile.
than the blue line, which converges slowly but nicely to the target. Any policy maker would prefer inflation as shown by the blue line over inflation as shown by the red line. However, $msd(\cdot)$ does not differentiate between these lines (while $v(\cdot)$ does).

There are other examples one can use to illustrate the differences between the measures. Imagine for example inflation staying constant for the first half of a time span at one percentage point below the target and then changing once and staying constant at one percentage point above the target. $msd(\cdot)$ does not distinguish between this very stable series and a series which randomly jumps back and forth between one percentage point below and one above the target (being at either value half of the time). $v(\cdot)$ distinguishes between these time series. $v(\cdot)$ is also not a perfect measure, however. For example if one were to compare inflation represented by two horizontal lines of which one is close to the target while the other is relatively far from the target, $v(\cdot)$ does not distinguish between these lines, while $msd(\cdot)$ does.

From a practitioner’s or policy maker’s point of view, which measure to use can thus depend on what kind of dynamics are present. For example if there are a lot of inflation time series which are relatively constant on one side of the target while some of these observations are close and some far from the target, $msd(\cdot)$ looks like a better measure. If one sees both erratic behavior or oscillations partly below and partly above the target and slow convergence, $v(\cdot)$ is the better measure. The latter case is exactly what we observe in the experiment. Inflation mainly oscillates around the target with mean values close to the target with some observations converging gradually to the target. From this point of view $v(\cdot)$ is clearly to be preferred.
D Appendix: Instructions in the Experiment

Subjects in the experiment received the following instructions (as subjects only received qualitative information on the model governing the experimental economy the instructions are the same for both treatments):

Instructions

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

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

General information and experimental economy

All participants will be randomly divided into groups of six people. The group composition will not change during the experiment. You and all other participants will take the roles of statistical research bureaus making predictions of inflation and the so-called “output gap”. The experiment consists of 50 periods in total. In each period you will be asked to predict inflation and output gap for the next period. The economy you are participating in is described by three variables: inflation \(\pi_t\), output gap \(y_t\), and interest rate \(i_t\). The subscript \(t\) indicates the period the experiment is in. In total there are 50 periods, so \(t\) increases during the experiment from 1 to 50.

Inflation

Inflation measures the percentage change in the price level of the economy. In each period, inflation depends on inflation predictions of the statistical research bureaus in the economy (a group of six participants in this experiment), on actual output gap and on a random term. There is a positive relation between the actual inflation and both inflation predictions and actual output gap. This means for example that if the inflation predictions of the research bureaus increase, then actual inflation will also increase (everything else equal). In economies similar to this one, inflation has historically been between \(-5\%\) and \(10\%\).
Output gap

The output gap measures the percentage difference between the Gross Domestic Product (GDP) and the natural GDP. The GDP is the value of all goods produced during a period in the economy. The natural GDP is the value the total production would have if prices in the economy were fully flexible. If the output gap is positive (negative), the economy therefore produces more (less) than the natural GDP. In each period the output gap depends on inflation predictions and output gap predictions of the statistical bureaus, on the interest rate and on a random term. There is a positive relation between the output gap and inflation predictions and also between the output gap and output gap predictions. There is a negative relation between the output gap and the interest rate. In economies similar to this one, the output gap has historically been between $-5\%$ and $5\%$.

Interest Rate

The interest rate measures the price of borrowing money and is determined by the central bank. If the central bank wants to increase inflation or output gap it decreases the interest rate, if it wants to decrease inflation or output gap it increases the interest rate.

Prediction task

Your task in each period of the experiment is to predict inflation and output gap in the next period. When the experiment starts, you have to predict inflation and output gap for the first two periods, i.e. $\pi^e_1$ and $\pi^e_2$, and $y^e_1$ and $y^e_2$. The superscript $e$ indicates that these are predictions. When all participants have made their predictions for the first two periods, the actual inflation ($\pi_1$), the actual output gap ($y_1$) and the interest rate ($i_1$) for period 1 are announced. Then period 2 of the experiment begins. In period 2 you make inflation and output gap predictions for period 3 ($\pi^e_3$ and $y^e_3$). When all participants have made their predictions for period 3, inflation ($\pi_2$), output gap ($y_2$), and interest rate ($i_2$) for period 2 are announced. This process repeats itself for 50 periods.

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

- Values of actual inflation, output gap and interest rate up to period $t - 1$;
- Your predictions up to period $t$;
- Your prediction scores up to period $t - 1$.

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

<table>
<thead>
<tr>
<th>Prediction error</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>100</td>
<td>50</td>
<td>33.33</td>
<td>25</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

**Example:** If (for a certain period) you predict an inflation of 2%, and the actual inflation turns out to be 3%, then you make an absolute error of $3\% - 2\% = 1\%$. Therefore you get a prediction score of 50. If you predict an inflation of 1%, and the actual inflation turns out to be negative 2% (i.e. $-2\%$), you make a prediction error of $1\% - (-2\%) = 3\%$. Then you get a prediction score of 25. For a perfect prediction, with a prediction error of zero, you get a prediction score of 100. The figure below shows the relation between your prediction score (vertical axis) and your prediction error (horizontal axis). Points in the graph correspond to the prediction scores in the previous table.

[Figure 25 appears here in the experimental instructions.]

At the end of the experiment, you will have two total scores, one for inflation predictions and one for output gap predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for inflation and output gap predictions. **When the experiment has ended, one of the two total scores will be randomly selected for payment.**

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

**Computer interface**

The computer interface will be mainly self-explanatory. The top right part of the screen will show you all of the information available up to the period that you are in (in period $t$, i.e. when you are asked to make your prediction for period $t+1$, this will be actual inflation, output gap, and interest rate until period $t-1$, your predictions until period
$t$, and the prediction scores arising from your predictions until period $t-1$ for both inflation ($I$) and output gap ($O$)). The top left part of the screen will show you the information on inflation and output gap in graphs. The axis of a graph shows values in percentage points (i.e. 3 corresponds to 3%). Note that the values on the vertical axes may change during the experiment and that they are different between the two graphs – the values will be such that it is comfortable for you to read the graphs.

In the bottom left part of the screen you will be asked to enter your predictions. When submitting your prediction, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 2.5% type “2.5”; for a prediction of −1.75% type “–1.75”. The sum of the prediction scores over the different periods are shown in the bottom right of the screen, separately for your inflation and output gap predictions.

At the bottom of the screen there is a status bar telling you when you can enter your predictions and when you have to wait for other participants.

Figure 25: Relation score and forecast error (not labeled in the instructions)
E  Appendix (for Online Publication): Graphs of the Experimental Data by Group and Screenshot

E.1  Realizations and Forecasts of Inflation and Output Gap in All Groups

Figures 26 to 34 show the realizations and forecasts of inflation and output gap. Each graph corresponds to one group of six people (one experimental economy). The thick black line shows the realization of inflation, the thin dashed black lines show the inflation forecasts of the six individuals in the group. The thick gray line shows the realization of the output gap and the thin dashed gray lines show the output gap forecasts of all individuals in a group. On the horizontal axis are the periods (from 1 to 50), on the vertical axis are the values of inflation and output gap in percent (the numbers on the vertical axis reach from $-3$ to 8). The upper red line corresponds to the steady state value of inflation ($\bar{\pi} = 3.5$), the lower red line corresponds to the steady state value of the output gap ($\bar{y} = 0.1166667$). Figures 26 to 29 show all groups of treatment $T_1$, Figures 30 to 33 show the groups of treatment $T_2$. Figure 34 shows the two groups (from $T_2$) that have been excluded from the analysis as explained in Footnote 8.

E.2  Screenshot

Figure 35 shows a screenshot (a larger version of the screenshot already used in Figure 4).
Figure 26: Realizations and forecasts of inflation and output gap ($T_{1}$, groups 1 – 6)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 27: Realizations and forecasts of inflation and output gap ($T_1$, groups 7 – 12)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 28: Realizations and forecasts of inflation and output gap (T1, groups 13 – 18)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 29: Realizations and forecasts of inflation and output gap ($T_1$, groups 19 – 21)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 30: Realizations and forecasts of inflation and output gap ($T_2$, groups 1 – 6)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 31: Realizations and forecasts of inflation and output gap ($T_2$, groups 7 – 12)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 32: Realizations and forecasts of inflation and output gap ($T^2$, groups 13 – 18)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 33: Realizations and forecasts of inflation and output gap ($T^2$, groups 19 – 22)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 34: Realizations and forecasts of inflation and output gap (excluded groups)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.
Figure 35: Screenshot

Information table

<table>
<thead>
<tr>
<th>Period</th>
<th>Inflation</th>
<th>Inflation Forecast</th>
<th>Output Gap</th>
<th>Output Gap Forecast</th>
<th>Interest Rate</th>
<th>Your Score (I)</th>
<th>Your Score (O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4.09</td>
<td>3.62</td>
<td>0.16</td>
<td>0.80</td>
<td>5.05</td>
<td>49.83</td>
<td>69.23</td>
</tr>
<tr>
<td>9</td>
<td>3.00</td>
<td>2.89</td>
<td>0.11</td>
<td>1.40</td>
<td>3.52</td>
<td>83.17</td>
<td>43.62</td>
</tr>
<tr>
<td>8</td>
<td>2.94</td>
<td>2.59</td>
<td>0.55</td>
<td>1.00</td>
<td>3.65</td>
<td>69.58</td>
<td>69.00</td>
</tr>
<tr>
<td>7</td>
<td>2.61</td>
<td>1.99</td>
<td>0.46</td>
<td>0.00</td>
<td>3.11</td>
<td>58.40</td>
<td>68.60</td>
</tr>
<tr>
<td>6</td>
<td>1.89</td>
<td>2.19</td>
<td>0.64</td>
<td>0.00</td>
<td>1.83</td>
<td>82.86</td>
<td>95.13</td>
</tr>
<tr>
<td>5</td>
<td>2.09</td>
<td>2.69</td>
<td>0.65</td>
<td>0.00</td>
<td>2.13</td>
<td>66.34</td>
<td>57.00</td>
</tr>
<tr>
<td>4</td>
<td>2.67</td>
<td>3.09</td>
<td>0.32</td>
<td>0.00</td>
<td>3.13</td>
<td>75.21</td>
<td>75.67</td>
</tr>
<tr>
<td>3</td>
<td>2.88</td>
<td>5.19</td>
<td>-3.26</td>
<td>0.20</td>
<td>3.14</td>
<td>31.67</td>
<td>66.91</td>
</tr>
<tr>
<td>2</td>
<td>4.07</td>
<td>5.09</td>
<td>-3.59</td>
<td>-1.00</td>
<td>5.98</td>
<td>80.62</td>
<td>70.90</td>
</tr>
</tbody>
</table>

Forecast Submission

You are now in period 10.

Enter your forecast for inflation in period 11: 

Enter your forecast for the output gap in period 11: 

Submit

Summary Information

Your total score for inflation is 605.09

Your total score for output gap is 617.14

Please submit your forecast.
F Appendix (for Online Publication): Additional Graphs and Data Analysis

F.1 Inflation in the Experiment

Figure 36 shows the empirical cumulative distribution functions of price instability when employing different measures. The first graph shows the volatility measure based on the absolute deviation, \( ad(\pi) = \frac{1}{T} \sum_{t=2}^{T} |\pi_t - \pi_{t-1}|. \) The second graph shows the means squared deviation from the target, \( msd(\pi) = \frac{1}{T} \sum_{t=2}^{T} (\pi_t - \bar{\pi})^2. \) and the third graph shows the standard deviation, \( sd(\pi) = \frac{1}{T} \sum_{t=2}^{T} (\pi_t - \pi^{av})^2. \) The average values for \( ad \) are 0.304 in \( T1 \) and 0.188 in \( T2. \) For \( msd, \) the values are 0.402 in \( T1 \) and 0.317 in \( T2, \) and for \( sd \) 0.510 in \( T1 \) and 0.419 in \( T2. \)
Figure 36: Empirical distribution functions of various measures

Notes: This graph shows the ECDFs for three different measures. From top to bottom: The volatility measure based on the absolute deviation, the mean squared deviation from target, and the standard deviation. For each value on the horizontal axis, the fraction of observations with the respective measure less or equal to this value (i.e. the ECDF) is shown on the vertical axis, separately for \( T_1 \) and \( T_2 \).