



**UvA-DARE (Digital Academic Repository)**

**Bounded rationality and heterogeneous expectations in macroeconomics**

Massaro, D.

[Link to publication](#)

*Citation for published version (APA):*

Massaro, D. (2012). *Bounded rationality and heterogeneous expectations in macroeconomics*. Amsterdam: Thela Thesis.

**General rights**

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

**Disclaimer/Complaints regulations**

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

# Chapter 3

## A Micro-Foundation of Heterogenous Expectations in New Keynesian Models

### 3.1 Introduction

Over the past decade the New Keynesian model has become increasingly popular in the analysis of monetary policy. This model is built under the hypothesis of rational expectations (RE) and assumes a representative agent structure. Although adaptive learning has become increasingly important as an alternative approach to modeling private sector expectations, most of these models still assume a representative agent who is learning about the economy (see e.g. Evans and Honkapohja (2001) and Sargent (1999) for extensive overviews). Carroll (2003), Mankiw, Reis, and Wolfers (2003), Branch (2004), Pfajfar and Santoro (2010) and Pfajfar (2008) recently provided empirical evidence in support of heterogeneous expectations using survey data on inflation expectations, while Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Adam (2007), Pfajfar and Zakelj (2010), Assenza, Heemeijer, Hommes, and Massaro (2011), and Hommes (2011) find evidence for heterogeneity in learning to forecast laboratory experiments with human subjects. Some recent examples of models with heterogeneous expectations in macroeconomics include Brock and de Fontnouvelle (2000), Evans and Honkapohja (2003,

2006), Branch and McGough (2009, 2010), Berardi (2007), Tuinstra and Wagener (2007), Brazier, Harrison, King, and Yates (2008), Anufriev, Assenza, Hommes, and Massaro (2008), and De Grauwe (2010). See also the works of Delli Gatti, Di Guilmi, Gaffeo, Giulioni, Gallegati, and Palestrini (2005), Agliari, Delli Gatti, Gallegati, and Lenci (2006), and Gaffeo, Delli Gatti, Desiderio, and Gallegati (2008) on heterogeneous agents in macroeconomics, and den Haan (2010) for a comparison of different algorithms to solve models with heterogeneous agents, idiosyncratic and aggregate risk.

In the light of the empirical evidence, the primary interest of this paper is to incorporate heterogeneous beliefs into the micro-foundations of the New Keynesian framework. The first contribution of our work is thus the development of a micro-founded framework for monetary policy analysis consistent with heterogeneous, possibly boundedly rational, expectations.

In our model agents solve infinite horizon decision problems<sup>1</sup>. The RE hypothesis requires that agents make fully optimal decisions given their beliefs and that agents know the true equilibrium distribution of variables that are beyond their control. Achieving the standard rationality requirements of RE models is especially difficult in a setting with heterogeneous agents. Individuals need to gather and process a substantial amount of information about the economy, including details about other agents in the market and their expectations, in order to derive the objective probability distributions of aggregate variables.

Starting from this observation, we assume that agents need to pay information gathering and processing costs<sup>2</sup> in order to achieve RE. As an alternative, individuals can use simple and freely available prediction rules (heuristics) to forecast aggregate variables. Our modeling approach slightly departs from the standard RE benchmark. In fact we assume costly information processing without excluding the RE predictor from the set of forecasting rules available to the agents in the economy. Moreover, while we assume cognitive limitations of individual under-

---

<sup>1</sup>Models with this kind of approach have been studied by Marcet and Sargent (1989) and Sargent (1993) among others.

<sup>2</sup>The idea of costly information processing is shared with the literature on “rational inattention” (see Sims (2003) among others).

standing<sup>3</sup>, we fully maintain the assumption that individual choices are made optimally given subjective, possibly non-rational, forecasts. The decision whether to pay the costs for the RE predictor will depend on the dynamic trade off between these costs and the forecasting performance of the heuristics along the lines of Brock and Hommes (1997). Agents' beliefs determine aggregate outcomes and are subsequently updated based upon recent performances when new public information becomes available in a sort of Bayesian updating mechanism. Co-evolution of subjective expectations with observed macroeconomic outcomes emerges due to the ongoing evaluation of predictors in a dynamic feedback system (Diks and van der Weide (2005)).

As underscored by Preston (2005), when the micro-foundations underpinning the New Keynesian model are solved under the non-rational expectations assumption, the predicted aggregate dynamics depend on long horizon forecasts. Hence, in making current decisions about spending and pricing of their output, agents take into account forecasts of macroeconomic conditions over an infinite horizon. This is a direct consequence of the fact that individuals are assumed to only have knowledge of their own objectives and of the constraints they face, and they do not have a complete economic model of determination of aggregate variables. We derive aggregate demand and supply equations consistent with heterogeneous expectations by explicitly aggregating individual decision rules. The resulting reduced form model is analytically tractable and encompasses the representative rational agent benchmark as a special case. In fact, when RE agents are present in the economy, it is possible to reduce the aggregate equations depending on long horizon forecasts to a system where only one-period ahead forecasts matter by using the law of iterated expectations. When all agents have RE, the model with forecasts over the infinite horizon reduces to the standard New Keynesian model.

Within a general framework of heterogeneous expectations, we test the desirability of standard policy recommendations. We find that bounded rationality represents an important source of instability and indeterminacy. In fact, monetary policy rules that lead to determinacy in a world with a representative rational agent may destabilize the economy even when only a small

---

<sup>3</sup>Cognitive limitations of individuals have been abundantly documented by psychologists and brain scientists. For recent surveys see Kahneman and Thaler (2006) and Della Vigna (2007).

fraction of boundedly rational agents is added to the system. Our results thus confirm the concerns for private sector expectations of policy makers such as Bernanke (2004) and show the importance of taking bounded rationality into account when designing monetary policies.

Closely related to our results is the parallel paper by Branch and McGough (2009) who also introduce heterogeneous expectations in a New Keynesian framework. Differently from our work, Branch and McGough (2009) start from the assumption that agents with subjective (non-rational) expectations choose optimal plans that satisfy the associated Euler equations instead of looking at the perceived lifetime intertemporal budget constraint<sup>4</sup>. The consequence of this behavioral assumption is that in Branch and McGough (2009) only one-period ahead forecasts matter for aggregate dynamics. This represents the main difference with our framework where individual decision rules as well as aggregate equations depend on long horizon forecasts.

The paper is organized as follows. Section 3.2 briefly recalls the standard New Keynesian model with a representative rational agent. The general framework consistent with heterogeneous expectations is derived in Section 3.3. Section 3.4 presents an application to monetary policy by considering determinacy issues in an economy with a continuum of boundedly rational beliefs together with perfectly rational expectations. Finally, Section 3.5 concludes.

## **3.2 The model with a representative (rational) agent**

In this section we briefly recall the standard New Keynesian model of output and inflation determination under RE<sup>5</sup>.

### **3.2.1 Households**

The preferences of the representative household are defined over a composite consumption good  $c_t$  and time devoted to market employment  $h_t$ . Households maximize the expected present

---

<sup>4</sup>See Preston (2005) and Evans, Honkapohja, and Mitra (2003) for a discussion about the two different modeling approaches.

<sup>5</sup>For a detailed presentation of the New Keynesian framework see Woodford (2003) and Galí (2008) among others.

discounted value of utility:

$$\max E_t \sum_{s=t}^{\infty} \delta^{s-t} \left( \frac{c_s^{1-\sigma}}{1-\sigma} - \chi \frac{h_s^{1+\gamma}}{1+\gamma} \right).$$

The composite consumption good consists of differentiated products produced by monopolistically competitive final good producers. There is a continuum of goods of measure 1, each of them produced by firms indexed by  $j \in [0, 1]$ . The composite consumption good that enters the households' utility function and the aggregate price index for consumption are the usual CES aggregators, defined as

$$c_t = \left( \int_0^1 c_{j,t}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \quad \text{and} \quad P_t = \left( \int_0^1 P_{j,t}^{1-\eta} dj \right)^{\frac{1}{1-\eta}}, \quad (3.1)$$

where the parameter  $\eta$  governs the price elasticity of demands for individual goods. The budget constraint of the household can be written in real terms as

$$c_t + b_t \leq w_t h_t + R_{t-1} \pi_t^{-1} b_{t-1} + d_t, \quad (3.2)$$

where  $b_t$  represents holdings of one-period bonds,  $w_t$  is the real wage,  $R_t$  is the (gross) nominal interest rate,  $\pi_t \equiv P_t/P_{t-1}$  is the inflation between period  $t$  and period  $t-1$  and  $d_t$  are dividends received from firms.

The first order conditions for the optimization problem are:

$$c_t^{-\sigma} = E_t (\delta R_t \pi_{t+1}^{-1} c_{t+1}^{-\sigma}) \quad (3.3)$$

$$\chi h_t^\gamma = w_t c_t^{-\sigma} \quad (3.4)$$

plus the budget constraint (3.2) holding as an equality and the transversality condition

$$\lim_{s \rightarrow \infty} \delta^{s-t} E_t c_s^{-\sigma} b_s = 0. \quad (3.5)$$

### 3.2.2 Firms

We assume a continuum of firms indexed by  $j \in [0, 1]$  producing differentiated goods. The production function is assumed to be linear in labor:

$$y_{j,t} = h_{j,t}. \quad (3.6)$$

The real profits of firm  $j$  are

$$\frac{P_{j,t}}{P_t} y_{j,t} - \frac{W_t}{P_t} h_{j,t} = \left( \frac{P_{j,t}}{P_t} - \frac{W_t}{P_t} \right) \left( \frac{P_{j,t}}{P_t} \right)^{-\eta} c_t,$$

where we substituted out for the demand for firm  $j$ 's output at time  $t$  constraint

$$y_{j,t} = c_{j,t} = \left( \frac{P_{j,t}}{P_t} \right)^{-\eta} c_t.$$

Assuming a Calvo staggered price setting, a firm able to reset its price is choosing  $P_{j,t}$  to solve

$$\max E_t \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( \frac{P_{j,t}}{P_s} - \frac{W_s}{P_s} \right) \left( \frac{P_{j,t}}{P_s} \right)^{-\eta} c_s,$$

where  $\omega$  is the degree of price stickiness and  $Q_s$  is the stochastic discount factor given by  $\delta^{s-t} (c_s/c_t)^{-\sigma}$ . Since all firms adjusting their price at time  $t$  are facing the same problem, they will set the same price,  $P_t^*$ . The first order condition for firms adjusting their price is

$$E_t \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( (1-\eta) \frac{1}{P_s} + \eta \frac{1}{P_t^*} w_s \right) \left( \frac{P_t^*}{P_s} \right)^{-\eta} c_s = 0. \quad (3.7)$$

Rearranging terms and using the definition of  $Q_s$  we can rewrite equation (3.7) as

$$\frac{P_t^*}{P_t} = \frac{\eta}{\eta-1} \frac{E_t \sum_{s=t}^{\infty} (\omega \delta)^{s-t} c_s^{-\sigma} w_s \left( \frac{P_s}{P_t} \right)^{\eta}}{E_t \sum_{s=t}^{\infty} (\omega \delta)^{s-t} c_s^{-\sigma} \left( \frac{P_s}{P_t} \right)^{\eta-1}}. \quad (3.8)$$

Using the price index definition (3.1) we can write the aggregate price level as

$$P_t = (\omega P_{t-1}^{1-\eta} + (1-\omega)(P_t^*)^{1-\eta})^{\frac{1}{1-\eta}}. \quad (3.9)$$

### 3.2.3 Government

We assume that the government issues no debt so that

$$b_t = 0, \quad \text{for all } t,$$

and that public expenditure is equal to zero. The market clearing condition thus requires

$$y_t = c_t.$$

Moreover, aggregating over firms, the total dividends transferred to the representative household are given by

$$d_t = y_t - w_t h_t.$$

### 3.2.4 Aggregate equations

In order to derive aggregate equations for output and inflation, we linearize the set of first order conditions around the steady state derived in Appendix 3.A. Log-linearizing the household's first order condition (3.3) and using the market clearing condition we get the expression

$$\hat{y}_t = E_t \hat{y}_{t+1} - \frac{1}{\sigma} (\hat{R}_t - E_t \hat{\pi}_{t+1}) \quad (3.10)$$

where  $\hat{x}_t \equiv \ln(x_t/x)$  denotes the log deviation of  $x_t$  from its steady state value  $x$ . Equation (3.10) is referred to as the *dynamic IS equation* (IS).



Log-linearizing equations (3.8) and (3.9) gives the expression for aggregate inflation

$$\widehat{\pi}_t = \delta E_t \widehat{\pi}_{t+1} + \widetilde{k} \widehat{w}_t, \quad (3.11)$$

where  $\widetilde{k} = \frac{(1-\omega)(1-\delta\omega)}{\omega}$ . Using the linearized version of the equilibrium condition (3.4) and the linearity of the production function<sup>6</sup> we can write  $\widehat{w}_t = \gamma \widehat{h}_t + \sigma \widehat{y}_t = (\sigma + \gamma) \widehat{y}_t$ , so that (3.11) becomes

$$\widehat{\pi}_t = \delta E_t \widehat{\pi}_{t+1} + k \widehat{y}_t, \quad (3.12)$$

where  $k = (\sigma + \gamma) \frac{(1-\omega)(1-\delta\omega)}{\omega}$ . Equation (3.12) is referred to as *New Keynesian Phillips Curve* (NKPC).

### 3.3 The model with heterogeneous expectations

We now depart from the benchmark model of Section 3.2 by introducing heterogeneous agents with *possibly* non-rational expectations. The framework is in fact very general, so that the sophisticated rational expectation predictor belongs to the class of available forecasting rules. In what follows we keep the assumption that the government issues no debt and public expenditure is equal to zero.

#### 3.3.1 Households

Consider an economy populated by a set of agents distributed uniformly along the unit interval with a total population mass of 1. That is, for  $i \in [0, 1]$  the probability distribution function is  $f(i) = 1$  and  $f(i) = 0$  elsewhere. The maximization problem of household  $i$  reads as

$$\max \widetilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( \frac{c_{i,s}^{1-\sigma}}{1-\sigma} - \chi \frac{h_{i,s}^{1+\gamma}}{1+\gamma} \right)$$

---

<sup>6</sup>Note that  $\widehat{y}_t = \widehat{h}_t$  up to a first order approximation in a neighborhood of the zero inflation steady state, see Galí (2008).

subject to

$$c_{i,t} + b_{i,t} \leq w_t h_{i,t} + R_{t-1} \pi_t^{-1} b_{i,t-1} + d_t,$$

where  $\tilde{E}_i$  denotes the subjective expectation of household  $i$ . The first order conditions for the problem are given by

$$\begin{aligned} c_{i,t}^{-\sigma} &= \tilde{E}_{i,t} (\delta R_t \pi_{t+1}^{-1} c_{i,t+1}^{-\sigma}) \\ \chi h_{i,t}^\gamma &= w_t c_{i,t}^{-\sigma} \end{aligned}$$

which can be log-linearized around the steady state derived in Appendix 3.A in order to get

$$\hat{c}_{i,t} = \tilde{E}_{i,t} \hat{c}_{i,t+1} - \sigma^{-1} (\hat{R}_t - \tilde{E}_{i,t} \hat{\pi}_{t+1}) \quad (3.13)$$

$$\hat{h}_{i,t} = \gamma^{-1} (\hat{w}_t - \sigma \hat{c}_{i,t}) \quad (3.14)$$

together with the budget constraint

$$\tilde{b}_{i,t} - \beta^{-1} \tilde{b}_{i,t-1} - b \beta^{-1} (\hat{R}_{t-1} - \hat{\pi}_t) - (1 - \eta^{-1}) (\hat{w}_t + \hat{h}_{i,t}) - \eta^{-1} \hat{d}_t + \hat{c}_{i,t} = 0, \quad (3.15)$$

where  $\hat{x}_t \equiv \ln(x_t/x)$  denotes the log deviation of  $x_t$  from its steady state value  $x$ , while  $\tilde{x}_t \equiv x_t - x$  is just the difference of variable  $x_t$  from its steady state  $x$ . Since the budget constraint has to hold as an equality in every period we can iterate it forward from period  $t$  on to get the intertemporal budget constraint defined as

$$\tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \hat{c}_{i,s} = \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left[ \delta^{-1} \tilde{b}_{i,t-1} + (1 - \eta^{-1}) (\hat{w}_s + \hat{h}_{i,s}) + \eta^{-1} \hat{d}_s \right], \quad (3.16)$$

where we used that  $b = 0$  and we imposed the No Ponzi constraint<sup>7</sup>

$$\lim_{s \rightarrow \infty} \tilde{E}_{i,t} \delta^{s-t} \tilde{b}_{i,s+1} = 0.$$

As shown in Appendix 3.B, after some algebraic manipulations we can derive the following consumption rule for agent  $i$ :

$$\begin{aligned} \hat{c}_{i,t} = & \frac{(1-\delta)\gamma}{\gamma + (1-\eta^{-1})\sigma} \left( \frac{1}{\delta(1-\delta)} \tilde{b}_{i,t-1} + \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \frac{(1-\eta^{-1})(1+\gamma)}{\gamma} \hat{w}_s + \eta^{-1} \hat{d}_s \right) \\ & - \frac{\delta}{\sigma} \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} (\hat{R}_s - \hat{\pi}_{s+1}). \end{aligned} \quad (3.17)$$

The individual consumption rule (3.17) can be interpreted in the spirit of the canonical consumption model. The first three terms reflect the basic insight that current consumption depends on the expected future discounted wealth, while the last term arises from the assumption of a time-varying real interest rate, and represents deviations from the equilibrium level  $R = \delta^{-1}$  due to either variations in the nominal interest rate or inflation.

### 3.3.2 Firms

The optimization problem for a firm able to reset its price reads

$$\max \tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( \frac{P_{j,t}}{P_s} - \frac{W_s}{P_s} \right) \left( \frac{P_{j,t}}{P_s} \right)^{-\eta} c_s,$$

where  $\tilde{E}_{j,t}$  denotes the subjective expectation of firm  $j$ . Since we are in an heterogeneous world and the optimization problem involves subjective expectations which can in principle differ among agents, we will have that in general different firms will set different prices. The average

---

<sup>7</sup>Even if expectations are not rational we are assuming that households are not allowed to believe that they can borrow (and consume) as much as they want and just pay off the interest payments by borrowing more.

price set by firms optimizing at time  $t$  will then be

$$P_t^* = \sum_j n_j P_{j,t},$$

where  $n_j$  denotes the fraction of firms setting price  $P_{j,t}$ . The first order condition for firm  $j$ 's problem is given by

$$\tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( (1-\eta) \frac{1}{P_s} + \eta \frac{1}{P_{j,t}} w_s \right) \left( \frac{P_{j,t}}{P_s} \right)^{-\eta} c_s = 0, \quad (3.18)$$

where again  $w_t = \frac{W_t}{P_t}$  denotes real marginal costs. Defining  $p_{j,t} \equiv \frac{P_{j,t}}{P_t}$  and  $\pi_{t,s} \equiv \frac{P_s}{P_t}$  we can rewrite (3.18) as

$$\begin{aligned} 0 &= \tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( (1-\eta) \frac{1}{P_s} P_{j,t} \frac{P_t}{P_t} + \eta P_{j,t} \frac{1}{P_{j,t}} w_s \right) \left( \frac{P_{j,t} P_t}{P_s P_t} \right)^{-\eta} c_s \\ &= \tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( (1-\eta) \frac{p_{j,t}}{\pi_{t,s}} + \eta w_s \right) \left( \frac{p_{j,t}}{\pi_{t,s}} \right)^{-\eta} c_s. \end{aligned} \quad (3.19)$$

Moreover we can rewrite the average price set by firms optimizing at time  $t$  as

$$p_t^* = \sum_j n_j p_{j,t},$$

so that the aggregate price level equation (3.9), multiplying both sides by  $P_t^{\eta-1}$ , can be written as

$$\left( \frac{P_t}{P_t} \right)^{1-\eta} = \omega \left( \frac{P_{t-1}}{P_t} \right)^{1-\eta} + (1-\omega) \left( \frac{P_t^*}{P_t} \right)^{1-\eta},$$

or, in terms of the inflation rate

$$1 = \omega \left( \frac{1}{\pi_t} \right)^{1-\eta} + (1-\omega) (p_t^*)^{1-\eta}. \quad (3.20)$$

Log-linearizing equation (3.20) gives

$$\widehat{\pi}_t = \frac{(1 - \omega)}{\omega} \widehat{p}_t^* \quad (3.21)$$

As shown in Appendix 3.C we can linearize (3.19) to get the following pricing equation:

$$\widehat{p}_{j,t} = (1 - \omega\delta) \widetilde{E}_{j,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \widehat{w}_s + \omega\delta \widetilde{E}_{j,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \widehat{\pi}_{s+1}. \quad (3.22)$$

Equation (3.22) shows that when a firm sets its price it must be concerned with future marginal costs and future inflation because it may be unable to adjust its price for several periods.

### 3.3.3 Aggregation of individual decision rules

We assume for simplicity that households are running firms (or CEOs are appointed by shareholders so they are discounting profits the same way shareholders would do) so expectations of households and firms will be the same. This means that indexes  $i$  and  $j$  will coincide. In this way we can also justify the fact that firms are using the stochastic multiplier  $Q_s = \delta^{s-t} (c_s/c_t)^{-\sigma}$  to discount future profits. Assume that there are different predictors available to agents in the economy. Agents using the same predictor will have the same expectation, so we can denote the fraction of agents adopting the forecasting rule  $h$  at time  $t$  as  $n_{h,t}$ .

Start from households consumption rule (3.17) and integrate over  $i$  to get

$$\begin{aligned} \widehat{c}_t = & \frac{(1 - \delta)\gamma}{\gamma + (1 - \eta^{-1})\sigma} \left( \frac{(1 - \eta^{-1})(1 + \gamma)}{\gamma} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} \widehat{w}_s + \frac{1}{\eta} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} \widehat{d}_s \right) \\ & - \frac{\delta}{\sigma} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} \left( \widehat{R}_s - \widehat{\pi}_{s+1} \right), \end{aligned} \quad (3.23)$$

where  $\widehat{c}_t = \int_0^1 \widehat{c}_{i,t} f(i) di$ ,  $\widetilde{E}_t = \int_0^1 \widetilde{E}_{i,t} f(i) di = \sum_{h=1}^H n_{h,t} \widetilde{E}_{h,t}$ , and we used that  $\int_0^1 \widetilde{b}_{i,t-1} f(i) di = 0$  by market clearing. We now want to derive an IS curve in terms of output and real interest rate. The optimality condition (3.14) can be rewritten in terms of the real wage, which is taken

as parametric by the individuals, as  $\widehat{w}_t = \gamma \widehat{h}_{i,t} + \sigma \widehat{c}_{i,t}$ . Aggregating over individuals we have that  $\widehat{w}_t = \gamma \widehat{h}_t + \sigma \widehat{c}_t$ . The market clearing condition implies  $\widehat{c}_t = \widehat{y}_t$ , so we can write  $\widehat{w}_t = \gamma \widehat{h}_t + \sigma \widehat{y}_t$ . In order to eliminate the term  $\widehat{h}_t$  we can use the production function  $\widehat{y}_t = \widehat{h}_t$ , so that

$$\widehat{w}_t = (\gamma + \sigma) \widehat{y}_t. \quad (3.24)$$

Moreover we can log-linearize the expression for total dividends  $d_t = y_t - w_t h_t$ , using relation (3.24) and steady state values, in order to get

$$\widehat{d}_t = (1 - (\eta - 1)(\gamma + \sigma)) \widehat{y}_t. \quad (3.25)$$

Substituting (3.24) and (3.25) into (3.23) and using  $\widehat{c}_t = \widehat{y}_t$  we get<sup>8</sup>

$$\begin{aligned} \widehat{y}_t = & \frac{(1 - \delta) \gamma}{\gamma + (1 - \eta^{-1}) \sigma} \left( \frac{(1 - \eta^{-1})(1 + \gamma)}{\gamma} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} (\gamma + \sigma) \widehat{y}_s + \right. \\ & \left. - \frac{\delta}{\sigma} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} (\widehat{R}_s - \widehat{\pi}_{s+1}) \right) \end{aligned}$$

Rearranging terms gives:

$$\begin{aligned} \widehat{y}_t = & \frac{(1 - \delta) \gamma}{\gamma + (1 - \eta^{-1}) \sigma} \left( \frac{(1 - \eta^{-1})(1 + \gamma)(\gamma + \sigma)}{\gamma} + \frac{1 - (\eta - 1)(\gamma + \sigma)}{\eta} \right) \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} \widehat{y}_s \\ & - \frac{\delta}{\sigma} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} (\widehat{R}_s - \widehat{\pi}_{s+1}), \end{aligned}$$

which can be simplified to obtain the *Heterogeneous Expectations IS* equation (HE-IS)

$$\widehat{y}_t = (1 - \delta) \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} \widehat{y}_s - \frac{\delta}{\sigma} \widetilde{E}_t \sum_{s=t}^{\infty} \delta^{s-t} (\widehat{R}_s - \widehat{\pi}_{s+1}). \quad (3.26)$$

Consider now firms' pricing rule (3.22). Using that  $\widehat{p}_t^* = \sum_j n_j \widehat{p}_{j,t}$  we obtain from (3.21)

---

<sup>8</sup>It is consistent to replace  $\widehat{w}_t = (\gamma + \sigma) \widehat{y}_t$  and  $\widehat{d}_t = (1 - (\eta - 1)(\gamma + \sigma)) \widehat{y}_t$  in equation (3.23) because we know that (ex post) these equalities will hold in each period  $t$ .

that

$$\hat{\pi}_t = \frac{(1-\omega)(1-\omega\delta)}{\omega} \tilde{E}_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{w}_s + \frac{(1-\omega)\omega\delta}{\omega} \tilde{E}_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_{s+1} \quad (3.27)$$

where again  $\tilde{E}_t = \int_0^1 \tilde{E}_{i,t} f(i) di = \sum_{h=1}^H n_{h,t} \tilde{E}_{h,t}$ . Using that  $\hat{w}_t = (\sigma + \gamma) \hat{y}_t$ , we can rewrite (3.27) in terms of output as

$$\hat{\pi}_t = \frac{(1-\omega)(1-\omega\delta)}{\omega} (\sigma + \gamma) \tilde{E}_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{y}_s + \frac{(1-\omega)\omega\delta}{\omega} \tilde{E}_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_{s+1}$$

to obtain the *Heterogeneous Expectations New Keynesian Phillips Curve* (HE-NKPC)

$$\hat{\pi}_t = k \tilde{E}_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{y}_s + (1-\omega)\delta \tilde{E}_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_{s+1}. \quad (3.28)$$

## The benchmark RE model as a special case

It is easy to see that, under the hypothesis of homogeneous rational agents, the HE-IS equation (3.26) and the HE-NKPC relation (3.28) can be reduced to the standard IS equation (3.10) and NKPC relation (3.12). Leading (3.28) one period ahead and taking rational expectations gives

$$\begin{aligned} E_t \hat{\pi}_{t+1} &= k E_t E_{t+1} \sum_{s=t+1}^{\infty} (\omega\delta)^{s-t-1} \hat{y}_s + (1-\omega)\delta E_t E_{t+1} \sum_{s=t+1}^{\infty} (\omega\delta)^{s-t-1} \hat{\pi}_{s+1} \\ E_t \hat{\pi}_{t+1} &= k E_t \sum_{s=t+1}^{\infty} (\omega\delta)^{s-t-1} \hat{y}_s + (1-\omega)\delta E_t \sum_{s=t+1}^{\infty} (\omega\delta)^{s-t-1} \hat{\pi}_{s+1}, \end{aligned}$$

where the second equality makes use of the law of iterated expectations. Rewriting (3.28) as

$$\hat{\pi}_t = k \hat{y}_t + (1-\omega)\delta E_t \hat{\pi}_{t+1} + k E_t \sum_{s=t+1}^{\infty} (\omega\delta)^{s-t} \hat{y}_s + (1-\omega)\delta E_t \sum_{s=t+1}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_{s+1}$$

then gives

$$\begin{aligned}\widehat{\pi}_t &= k\widehat{y}_t + (1 - \omega)\delta E_t\widehat{\pi}_{t+1} + \omega\delta E_t\widehat{\pi}_{t+1} \\ &= k\widehat{y}_t + \delta E_t\widehat{\pi}_{t+1}.\end{aligned}$$

For the IS equation we have that (3.26) can be rewritten as

$$\begin{aligned}\widehat{y}_t &= (1 - \delta)\widehat{y}_t + (1 - \delta)E_t \sum_{s=t+1}^{\infty} \delta^{s-t}\widehat{y}_s - \frac{\delta}{\sigma}E_t \sum_{s=t}^{\infty} \delta^{s-t} \left( \widehat{R}_s - \widehat{\pi}_{s+1} \right) \\ &= (1 - \delta)E_t \sum_{s=t+1}^{\infty} \delta^{s-t-1}\widehat{y}_s - \frac{1}{\sigma}E_t \sum_{s=t}^{\infty} \delta^{s-t} \left( \widehat{R}_s - \widehat{\pi}_{s+1} \right).\end{aligned}\quad (3.29)$$

Leading (3.29) one period ahead and taking rational expectations gives

$$E_t\widehat{y}_{t+1} = (1 - \delta)E_t \sum_{s=t+2}^{\infty} \delta^{s-t-2}\widehat{y}_s - \frac{1}{\sigma}E_t \sum_{s=t+1}^{\infty} \delta^{s-t-1} \left( \widehat{R}_s - \widehat{\pi}_{s+1} \right)$$

so that equation (3.29) can be rewritten as

$$\begin{aligned}\widehat{y}_t &= (1 - \delta)E_t\widehat{y}_{t+1} + (1 - \delta)E_t \sum_{s=t+2}^{\infty} \delta^{s-t-1}\widehat{y}_s - \frac{1}{\sigma}(\widehat{R}_t - E_t\widehat{\pi}_{t+1}) \\ &\quad - \frac{1}{\sigma}E_t \sum_{s=t+1}^{\infty} \delta^{s-t} \left( \widehat{R}_s - \widehat{\pi}_{s+1} \right) \\ &= (1 - \delta)E_t\widehat{y}_{t+1} + \delta E_t\widehat{y}_{t+1} - \frac{1}{\sigma}(\widehat{R}_t - E_t\widehat{\pi}_{t+1}) \\ &= E_t\widehat{y}_{t+1} - \frac{1}{\sigma}(\widehat{R}_t - E_t\widehat{\pi}_{t+1}).\end{aligned}$$

The analysis performed in this section shows that, when the microfoundations underpinning the New Keynesian model are solved under non rational heterogeneous expectations, the aggregate dynamics depend on long horizon forecasts<sup>9</sup> as can be seen in equations (3.26) and

---

<sup>9</sup>From a behavioral perspective it may seem a little awkward that boundedly rational agents are basing their decisions on long horizon forecasts. However this is a direct consequence of the fact that we are only departing from the benchmark model by relaxing the rationality assumption in the way agents form expectations, but we are keeping the assumption that agents behave optimally given their subjective beliefs.



(3.28). In fact, boundedly rational agents do not have a complete model of determination of aggregate variables. As noted in Preston (2005), neither the aggregate demand relation (3.26) nor the Phillips curve (3.28) can be simplified as in the rational expectations equilibrium analysis where, since expectations are taken with respect to the correct distribution of future endogenous variables, the law of iterated expectations holds at the aggregate level and therefore only one period ahead expectations matter for aggregate dynamics.

### 3.4 Monetary policy with heterogeneous expectations

In this section we use the model with heterogeneous expectations developed in Section 3.3 for monetary policy analysis. Contemporary policy discussions argued that a desirable interest rate rule has to involve feedback from endogenous variables such as inflation and/or real activity. Many authors considered simple interest rate rules with endogenous components of the form:

$$\widehat{R}_t = \phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t \quad (3.30)$$

and analyzed determinacy properties of the rational expectation equilibrium. Desirable monetary rules should avoid indeterminacy (i.e. multiple bounded equilibria) and sunspot fluctuations as emphasized in the original analysis of Sargent and Wallace (1975). Under rational expectations a necessary and sufficient condition for the equilibrium to be determinate is given by the *Taylor principle*:

$$k(\phi_\pi - 1) + (1 - \delta)\phi_y > 0, \quad (3.31)$$

stating that the monetary authority should respond to inflation and real activity by adjusting the nominal interest rate with “sufficient strength”<sup>10</sup>. Recent studies investigated the validity of such a policy recommendation when expectations depart from the rational benchmark. In the context of a New Keynesian monetary model Bullard and Mitra (2002) assume that agents do not initially have rational expectations, and that they instead form forecasts by using recursive

---

<sup>10</sup>See Woodford (2003) for a proof.

least squares. Using the *E-stability*<sup>11</sup> criterion they show that an interest rate rule that satisfies the Taylor principle induces learnability of the RE equilibrium. Preston (2005) studies the learnability of the RE equilibrium in a New Keynesian setting with long horizon forecasts. He shows that under least squares learning dynamics the Taylor principle (3.31) is necessary and sufficient for E-stability. However, most models studying the validity of classical policy recommendation in contexts that depart from the RE assumption assume a representative agent who is learning about the economy.

The focus of this section is to analyze the dynamical consequences of a policy regime as in (3.30) when agents have *heterogeneous beliefs*. The framework developed in the first part of the paper can now be applied to evaluate the desirability of the monetary policy rule (3.30) in the presence of heterogeneous expectations. In particular we want to investigate whether the standard advices to policy makers lead the dynamics to converge to the RE equilibrium in a world with heterogeneous expectations.

### **3.4.1 Specification of expectations and evolutionary dynamics**

In this section we characterize individuals' expectation schemes and describe the evolution of beliefs over time. There is ample empirical evidence documenting that private sector beliefs, when proxied by surveys, are characterized by heterogeneity. Carroll (2003) and Branch (2004) analyze the Michigan survey data on inflation expectations and find results pointing in the direction of an intermediate degree of rationality. Pfajfar and Santoro (2010) study the same survey and document the fact that agents in different percentiles of the survey seem to be associated with forecasting schemes characterized by different degrees of rationality. Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Adam (2007), Pfajfar and Zakelj (2010), and Assenza, Heemeijer, Hommes, and Massaro (2011) find evidence for pervasive heterogeneity of beliefs in learning to forecast experiments with human subjects. In particular, Pfajfar and Zakelj (2010) find a significant proportion of rational agents (around 40%) in a monetary policy

---

<sup>11</sup>For a detailed presentation of the E-stability concept and its relation with real time learning, see Evans and Honkapohja (2001)

experiment set up in the New Keynesian framework.

Building on the empirical evidence mentioned above, we assume that a fraction of agents are perfectly rational and the remainder are boundedly rational, using heuristics to forecast macroeconomic variables. We thus have in mind an economy in which some agents face cognitive problems in understanding and processing information and hence eventually make mistakes in forecasting macroeconomic variables while other agents, by paying some information gathering and processing costs  $C \geq 0$  per period, have rational expectations.

Before describing the full model with both perfectly rational and boundedly rational agents, we will characterize the dynamic feedback system in which macroeconomic variables and heterogeneous subjective expectations co-evolve over time along the lines of Brock, Hommes, and Wagener (2005) and Diks and van der Weide (2005). We assume that agents do not fully understand how macroeconomic variables are determined and hence have biased forecasts. One might think about an economy in which subject do not know the target of the monetary authority and have biased beliefs about it. Moreover, as a result of cognitive limitations, there are differences in the use of information and thus heterogeneity in individual forecasts. As argued in De Grauwe (2010), the assumption of a simple biased forecast can be viewed as a parsimonious representation of a world where agents do not know the underlying economic model and have a biased view about this model.

The point predictors used by the agents are represented in the belief space  $\Theta$  and parameterized by the belief parameter  $\theta$ . Each value  $\theta_i \in \Theta$  represents a strategy that fully characterizes the behavior of individual  $i$ . We will consider the simplest possible case of constant predictors<sup>12</sup>. Within this class of simple rules we allow for differences in the conclusions that agents draw when processing information, as well as biases and idiosyncracies. Therefore in every period we will have a distribution of point predictions<sup>13</sup>. This specification is quite general in

---

<sup>12</sup>Anufriev, Assenza, Hommes, and Massaro (2008) use a similar modeling approach in a simple model with inflation expectations while De Grauwe (2010) assumes that agents use simple constant heuristics to forecast output gap in the context of a behavioral DSGE model. In the experiments of Assenza, Heemeijer, Hommes, and Massaro (2011) individuals frequently employ constant forecasting rules.

<sup>13</sup>Anufriev, Assenza, Hommes, and Massaro (2008) use the concept of *Large Type Limit* (LTL) developed by Brock, Hommes, and Wagener (2005) to analyze inflation dynamics when the number of strategies available to the

the sense that the individual point predictions can be thought of as the outcome of any mental process or estimation technique. Alternatively one could think of agents selecting between various Bayesian forecasting models each differing by their very strong priors. From the modeler's point of view, the strategy  $\theta_{i,t}$  used by agent  $i$  at time  $t$  is a random variables distributed according to the probability density function  $\psi_t(\theta)$ . Given that agents have limited cognitive abilities, their forecasts will typically be biased<sup>14</sup>. In order to limit the wilderness of bounded rationality and avoid completely irrational behavior we have to introduce discipline in the selection of rules. We will achieve this discipline by subjecting the choice of heuristics to a fitness criterion, and by introducing a selection mechanism that allows agents to learn from their forecasting mistakes.

Predictors will, in fact, be evaluated according to an *evolutionary fitness* measure. Given the performance of predictors at time  $t - 1$ , denoted by  $U_{t-1}(\theta)$ , we assume as in Diks and van der Weide (2005) that the distribution of beliefs evolves over time as a function of past performances according to the *continuous choice model*:

$$\psi_t(\theta) = \frac{v(\theta)e^{\beta U_{t-1}(\theta)}}{Z_{t-1}}, \quad (3.32)$$

where  $Z_{t-1}$  is a normalization factor independent of  $\theta$  given by

$$Z_{t-1} = \int_{\Theta} v(\vartheta)e^{\beta U_{t-1}(\vartheta)} d\vartheta,$$

and  $v(\theta)$  is a *opportunity function* that can put different weights on different parts of the beliefs space. The parameter  $\beta$  refers to the *intensity of choice* and measures how sensitive agents are

---

agents tend to infinity. The concept of *Continuous Beliefs System* (CBS) developed by Diks and van der Weide (2002) and used in this paper is a generalization of the LTL concept (see Diks and van der Weide (2003) for a discussion).

<sup>14</sup>The point predictions of the heuristics can coincide with the perfect foresight point forecast because we did not make any restrictive assumption on the support of the distribution of beliefs.

to differences in performances. Notice that (3.32), which can be rewritten as

$$\psi_t(\theta) \propto v(\theta)e^{\beta U_{t-1}(\theta)}, \quad (3.33)$$

is a rule for updating the distribution of beliefs as new information becomes available similar to a Bayesian updating rule. In fact,  $v(\theta)$  plays a role similar to a prior, reflecting the a priori faith of individuals in parameters within certain regions of the parameter space, and  $\psi_t(\theta)$  to a posterior. The fitness measure enters the beliefs distribution through the term  $e^{\beta U_{t-1}(\theta)}$ , which plays a role similar to the likelihood in Bayesian statistics. In fact, a Bayesian updating rule in the usual form is recovered exactly when  $\beta = 1$  and the performance measure  $U_{t-1}(\theta)$  is the log-likelihood function of an econometric model, given the available observations. In order to simplify the analysis we will assume that the performance measure is given by the past squared forecast error<sup>15</sup>

$$U_{t-1}(\theta) = -(\theta - x_{t-1})^2, \quad (3.34)$$

for  $x \in \{y, \pi, R\}$ . Having observed and compared overall past performance, all agents subsequently adapt their beliefs. The distribution of beliefs at time  $t$  is thus given by means of the continuous choice model (3.32). Co-evolution of the distribution of beliefs with the observed aggregate variables thus emerges through the ongoing evaluation of predictors. As in Diks and van der Weide (2005) we assume a constant opportunity function  $v(\theta) = 1$ , meaning that agents assign the same initial weight to all parameter values. Since the utility function is a quadratic function in the belief parameter  $\theta$ , it follows that, for all  $t$ , the distribution of beliefs is described by a normal distribution:

$$\psi_t(\theta) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{1}{2}\left(\frac{\theta - \mu_t}{\sigma_t}\right)^2\right). \quad (3.35)$$

---

<sup>15</sup>Branch (2004) finds empirical evidence for dynamic switching depending on the squared errors of the predictors in survey data on individuals' expectations, while Pfajfar and Zakelj (2010) and Hommes (2011) find empirical evidence for dynamic switching depending on the squared errors of the predictors in experimental data on individuals' expectations.

Substituting the specified performance measure  $U_{t-1}(\theta) = -(\theta - x_{t-1})^2$  and opportunity function  $v(\theta) = 1$  in (3.33) we get

$$\psi_t(\theta) \propto \exp(-\beta(\theta - x_{t-1})^2). \quad (3.36)$$

By comparing the exponents in equations (3.35) and (3.36), the mean  $\mu_t$  and the variance  $\sigma_t^2$  can be seen to evolve according to

$$\mu_t = x_{t-1} \quad (3.37)$$

$$\sigma_t^2 = \frac{1}{2\beta}, \quad (3.38)$$

while the remaining terms independent of  $\theta$  are accounted for by the normalization factor  $Z_{t-1}$ . The dynamics of  $\mu_t$  and  $\sigma_t^2$  fully characterize the evolution of  $\psi_t(\theta)$  over time. Dividing the continuum of agents interval  $[0,1]$  in  $n$  parts and taking the limit for  $n \rightarrow \infty$  we can write

$$\int_0^1 \theta_{i,t} f(i) di = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \theta_{j,t}.$$

In the absence of dependence among agents, the law of large number applies, and it follows that the average belief will converge to  $\mu_t$ , that is

$$\int_0^1 \theta_{i,t} f(i) di = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \theta_{j,t} = \mathbb{E}[\theta_{i,t}] = \int_{\Theta} \vartheta \psi_t(\vartheta) d\vartheta = \mu_t. \quad (3.39)$$

## Rational versus biased beliefs

We now turn to the description of the heterogeneous expectations model with rational and biased beliefs. Introducing rational agents in the economy we can rewrite the aggregate equations

(3.26) and (3.28) as

$$\begin{aligned}
\hat{y}_t &= n_{RE,t} E_t \sum_{s=t}^{\infty} \delta^{s-t} \left( (1-\delta)\hat{y}_s - \frac{\delta}{\sigma}(\hat{R}_s - \hat{\pi}_{s+1}) \right) \\
&\quad + \int_{n_{RE,t}}^1 \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( (1-\delta)\hat{y}_s - \frac{\delta}{\sigma}(\hat{R}_s - \hat{\pi}_{s+1}) \right) f(i) di \\
\hat{\pi}_t &= n_{RE,t} E_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} (k\hat{y}_s + (1-\omega)\delta\hat{\pi}_{s+1}) \\
&\quad + \int_{n_{RE,t}}^1 \tilde{E}_{i,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} (k\hat{y}_s + (1-\omega)\delta\hat{\pi}_{s+1}) f(i) di, \tag{3.40}
\end{aligned}$$

where  $n_{RE,t}$  denotes the fraction of rational agents at time  $t$  and  $E_t$  is the rational expectation operator.

We follow Kreps (1998) and Sargent (1999) in assuming that agents solve an anticipated utility problem, i.e. when agents solve their optimization problem they hold their expectation operator fixed and assume that it remains fixed for all future periods. Using this assumption we can simplify boundedly rational individual forecasts over the long horizon as

$$\tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \hat{x}_s = \frac{1}{1-\delta} \theta_{i,x,t}, \tag{3.41}$$

where  $\theta_{i,x,t}$  denotes the biased forecasts of agent  $i$  in period  $t$  for  $x \in \{y, \pi, R\}$ . Using the results in (3.37), (3.38) and (3.39) we have that the average forecast of boundedly rational biased agents is given by

$$\int_{n_{RE,t}}^1 \theta_{i,x,t} f(i) di = (1 - n_{RE,t}) \mu_{x,t} = (1 - n_{RE,t}) \hat{x}_{t-1}, \tag{3.42}$$

for  $x \in \{y, \pi, R\}$ . Therefore, using the results (3.41) and (3.42), we can rewrite system (3.40)

as

$$\begin{aligned}\widehat{y}_t &= n_{RE,t} E_t \sum_{s=t}^{\infty} \delta^{s-t} \left( (1 - \delta) \widehat{y}_s - \frac{\delta}{\sigma} (\widehat{R}_s - \widehat{\pi}_{s+1}) \right) + (1 - n_{RE,t}) V_y(\widehat{y}_{t-1}, \widehat{R}_t, \widehat{R}_{t-1}, \widehat{\pi}_{t-1}) \\ \widehat{\pi}_t &= n_{RE,t} E_t \sum_{s=t}^{\infty} (\omega \delta)^{s-t} (k \widehat{y}_s + (1 - \omega) \delta \widehat{\pi}_{s+1}) + (1 - n_{RE,t}) V_{\pi}(\widehat{y}_{t-1}, \widehat{\pi}_{t-1}),\end{aligned}\quad (3.43)$$

where  $V_x(\cdot)$  denotes a function linear in its arguments, for  $x \in \{y, \pi\}$ , described in equations (3.56) and (3.57) in Appendix 3.D.

At the beginning of each period agents have to decide whether to pay an information gathering and processing cost  $C \geq 0$  or use simple heuristics in order to forecast macroeconomic variables. We use the *discrete choice model* of Brock and Hommes (1997) to model this decision between the costly rational predictor or freely use a simple biased rule. We now recall the discrete choice model in general terms and then we apply the main results to our framework. Assume that in general agents can form expectations choosing from  $H$  different forecasting strategies. The fraction of individuals using the forecasting rule  $h$  at time  $t$  is denoted by  $n_{h,t}$ . With dynamic predictor selection the fractions are updated in every period according to the well known multinomial logit law of motion (see Manski and McFadden (1981) for details)

$$n_{h,t} = \frac{e^{\beta U_{h,t-1}}}{\sum_{h=1}^H e^{\beta U_{h,t-1}}}, \quad (3.44)$$

where  $U_{h,t-1}$  is the fitness metric of predictor  $h$  at time  $t - 1$ . Note that the higher the past performance of a forecasting rule  $h$ , the higher the probability that an agent will select strategy  $h$ . The parameter  $\beta$  refers to the intensity of choice and it reflects the sensitivity of the mass of agents to selecting the optimal prediction strategy. The case  $\beta = 0$  corresponds to the situation in which differences in fitness can not be observed, so agents do not switch between strategies and all fractions are constant and equal to  $1/H$ . The case  $\beta = \infty$  corresponds to the “neoclassical” limit in which the fitness can be observed perfectly and in every period all agents choose the best predictor.



In our setup the discrete logit model is used to model the choice between two different *classes* of predictors, namely the RE predictor and the heuristics. Assuming that the fitness measure is given by past squared forecast error minus the cost of the predictor, in order to decide whether to be rational or not, each individual is comparing the cost  $C$  (since the RE predictor corresponds to perfect foresight and thus has zero forecast error) with the average past forecast error of the freely available heuristics, summed over the three variables being forecast, given by<sup>16</sup>

$$\sum_x \left( \int (\theta_{x,t-1} - \hat{x}_{t-1})^2 \psi_{t-1}(\theta_x) d\theta_x \right) = \frac{3}{2\beta},$$

for  $x \in \{y, \pi, R\}$ . We thus have that the fraction  $n_{RE,t}$  is constant over time and given by

$$n_{RE} = \frac{e^{-\beta_d C}}{e^{-\beta_d C} + e^{-\beta_d \frac{3}{2\beta_c}}}, \quad (3.45)$$

where  $\beta_d$  and  $\beta_c$  denote respectively the intensities of choice of the discrete and continuous choice model.

Using the results derived before and closing the model with the interest rate rule (3.30) we can rewrite system (3.43) in the standard matrix form as<sup>17</sup>

$$\begin{bmatrix} E_t \hat{y}_{t+1} \\ E_t \hat{\pi}_{t+1} \\ \widehat{Y1}_{t+1} \\ \widehat{\Pi1}_{t+1} \end{bmatrix} = \begin{bmatrix} \gamma_{yy} & \gamma_{y\pi} & \gamma_{yY1} & \gamma_{y\Pi1} \\ \gamma_{\pi y} & \gamma_{\pi\pi} & \gamma_{\pi Y1} & \gamma_{\pi\Pi1} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{y}_t \\ \hat{\pi}_t \\ \widehat{Y1}_t \\ \widehat{\Pi1}_t \end{bmatrix}, \quad (3.46)$$

where  $\widehat{Y1}_t = \hat{y}_{t-1}$  and  $\widehat{\Pi1}_t = \hat{\pi}_{t-1}$ . We are now ready to investigate the validity of the standard monetary policy recommendations in the context of a New Keynesian model with a rich variety of forecasting rules, including the rational expectations predictor.

<sup>16</sup>Note that we are assuming that the intensity of choice governing the evolution of the distributions  $\psi_t(\theta_y)$ ,  $\psi_t(\theta_\pi)$ , and  $\psi_t(\theta_R)$  is the same. Relaxing this assumption does not alter our qualitative results.

<sup>17</sup>Description of the coefficients is given in Appendix 3.D.

### 3.4.2 Numerical analysis of determinacy

Model (3.46) has the form of a rational expectations model with predetermined variables. Techniques for analyzing the determinacy properties of a linear model under rational expectations are well known (see, for example, Blanchard and Kahn (1980)). Define the transition matrix in (3.46) as

$$\mathbb{M} = \begin{bmatrix} \gamma_{yy} & \gamma_{y\pi} & \gamma_{yY1} & \gamma_{y\Pi1} \\ \gamma_{\pi y} & \gamma_{\pi\pi} & \gamma_{\pi Y1} & \gamma_{\pi\Pi1} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

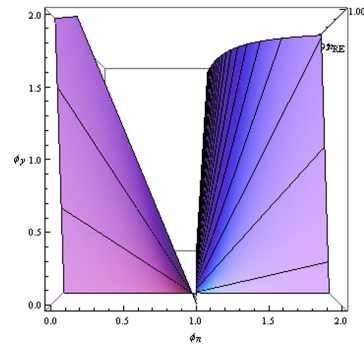
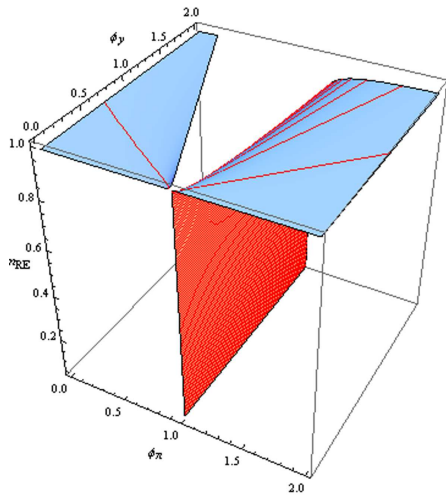
The determinacy properties of the model depend on the magnitude of the eigenvalues of  $\mathbb{M}$ . In our model with heterogeneous agents we have two predetermined variables, so that determinacy obtains when two eigenvalues are inside the unit circle. Fewer eigenvalues inside the unit circle imply explosiveness and more imply indeterminacy. Given the large number of the model's parameters we will perform a numerical analysis of determinacy properties. The model's parameters could be estimated but such an estimation is beyond the scope of this simple monetary policy exercise. We will therefore use calibrated values for the structural parameters as in Woodford (2003), namely  $\delta = 0.99$ ,  $\sigma = 0.157$ ,  $k = 0.024$  and  $\omega = 0.66$ , and treat the policy parameters  $\phi_\pi$  and  $\phi_y$ , together with the fraction of rational agents  $n_{RE}$ , as bifurcation parameters<sup>18</sup>. We will follow Branch and McGough (2009) and use  $0 < \phi_\pi < 2$ ,  $0 < \phi_y < 2$  as the benchmark policy space.

Fig. 3.1 shows the bifurcation surfaces at which eigenvalues of  $\mathbb{M}$  cross the unit circle in the 3-dimensional  $(\phi_\pi, \phi_y, n_{RE})$  parameters space<sup>19</sup>. By looking at the graphs in Fig. 3.1 it is possible to study how the bifurcations loci shift in the benchmark policy space as a continuous function of the degree of rationality in the economy.

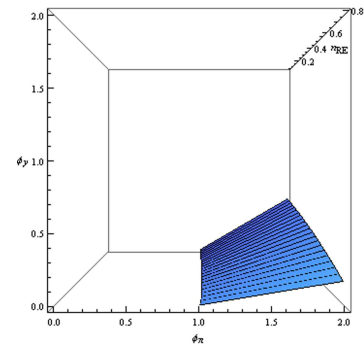
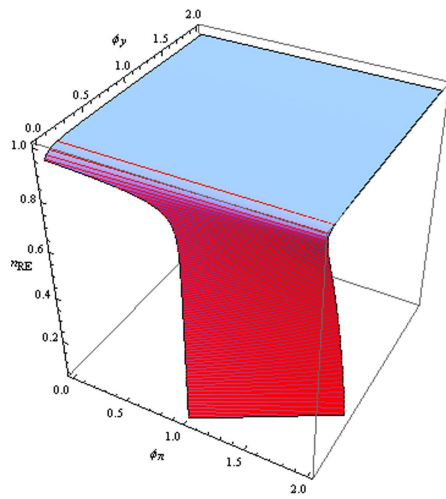
Fig. 3.1a refers to the *fold* bifurcation, associated with a qualitative change in the behavior

<sup>18</sup>The fundamental result of this section, namely that the presence of bounded rationality represents an important source of instability which may alter significantly the determinacy properties of the model, is robust across calibrations.

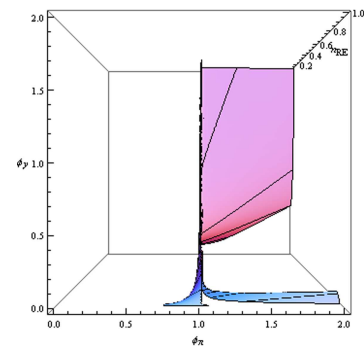
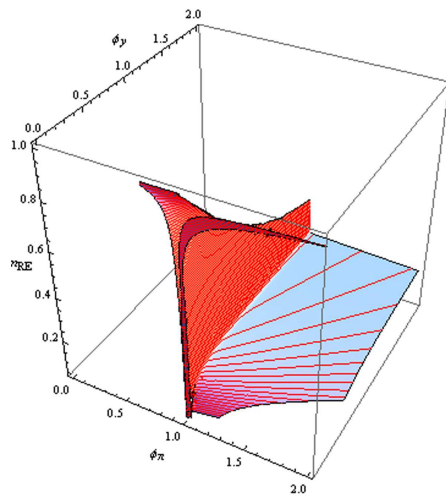
<sup>19</sup>See Kuznetsov (1995) for a detailed mathematical treatment of bifurcation theory.



(a) Fold



(b) Flip



(c) Neimark-Sacker

Figure 3.1: **Left panels:** Bifurcation surfaces. **Right panels:** Topviews.

of the economy due to the appearance of an eigenvalue  $\lambda_1 = 1$  of the matrix  $\mathbb{M}$ .

Fig. 3.1b refers to the *flip* bifurcation, associated with a qualitative change in the behavior of the economy due to the appearance of an eigenvalue  $\lambda_1 = -1$  of the matrix  $\mathbb{M}$ .

Fig. 3.1c refers to the *Neimark-Sacker* bifurcation, associated with a qualitative change in the behavior of the economy due to the appearance of complex eigenvalues  $\lambda_{1,2} = e^{\pm i\vartheta}$  of the matrix  $\mathbb{M}$ , where  $\vartheta$  is an arbitrary angle. Details about the computation of the bifurcation surfaces are provided in Appendix 3.E.

In order to draw conclusions about the determinacy properties of the model, one would need to overlap Figs. 3.1a, 3.1b, and 3.1c, and study the regions resulting from the intersections of the bifurcation surfaces. Even though this does not present particular problems from the computational point of view, it is easier to visualize the results in the 2-dimensional  $(\phi_\pi, \phi_y)$  parameter space for a given fraction of rational agents  $n_{RE}$ . The graphs in Fig. 3.1 are then useful to understand how the bifurcation curves shift in the 2-dimensional  $(\phi_\pi, \phi_y)$  space as more and more boundedly rational agents are introduced in the economy.

Fig. 3.2 shows how the determinacy properties of the model change when we allow for small departures from the representative rational agent benchmark. We present determinacy results for some empirically relevant cases. The choice of the fractions of rational agents in the economy reflects the findings of Galí and Gertler (1999) who estimated a reduced form NKPC with a fraction of forward looking rational firms and a fraction of firms with backward looking behavior and found a degree of rationality between 0.6 and 0.8, and the findings of Pfajfar and Zakelj (2010) who provided evidence of a degree of rationality around 0.4 in a monetary policy laboratory experiment.

Fig. 3.2a indicates the outcome under full rationality and it is consistent with the usual prescription of the RE literature: an interest rate rule satisfying the Taylor principle ensures determinacy. However, when we add a fraction of boundedly rational agents in the economy, the boundaries of the determinacy regions are sensibly altered. We indeed observe in Fig. 3.2b that when a small portion of non-rational agents is added to the economy, the region corresponding

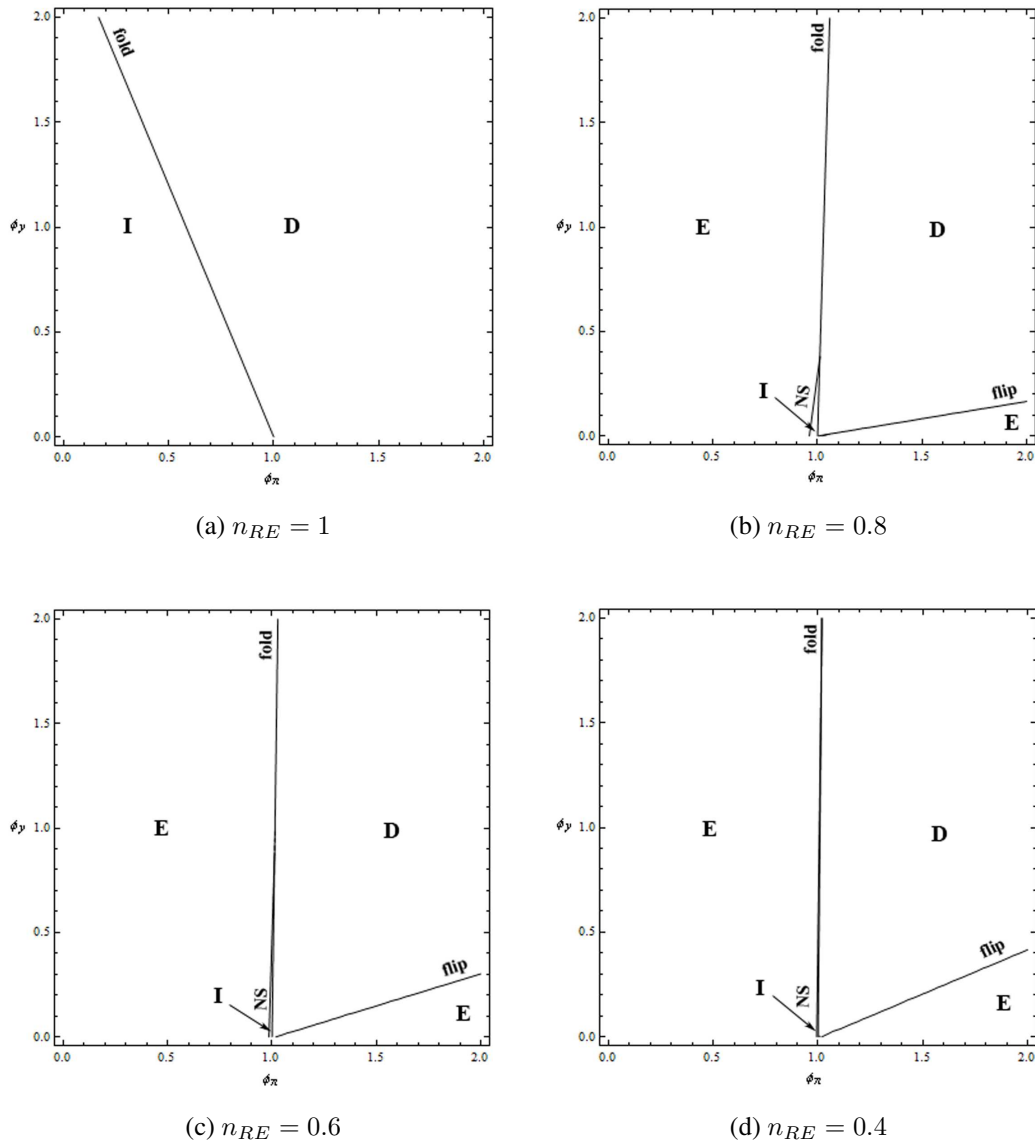


Figure 3.2: Determinacy properties for different values of the fraction  $n_{RE}$ .

**D** denotes determinacy, **I** denotes indeterminacy, and **E** denotes explosiveness.

The labels “fold”, “flip”, and “NS” correspond respectively to fold, flip, and Neimark-Sacker bifurcations.

**(a)**: RE benchmark,  $n_{RE} = 1$ , **(b)**: Galí and Gertler (1999) (upper bound),  $n_{RE} = 0.8$ , **(c)**: Galí and Gertler (1999) (lower bound),  $n_{RE} = 0.6$ , **(d)**: Pfajfar and Zakelj (2010),  $n_{RE} = 0.4$ .

to determinacy decreases in size, meaning that policy rules that obey the Taylor principle may not enforce determinacy. By looking at Fig. 3.1a we notice that the curve separating the regions of determinacy and indeterminacy in the benchmark RE case rotates counterclockwise and dis-

appears from the relevant policy space as the fraction of rational agents decreases. Moreover, a new fold bifurcation curve emerges and rotates counterclockwise as  $n_{RE}$  decreases. This curve corresponds to the curve labeled as “fold” in Fig. 3.2b. In the passage from Fig. 3.2a to Fig. 3.2b we notice that both a Neimark-Sacker and a flip bifurcation curve, respectively labeled as “N-S” and “flip”, emerge. These bifurcation curves separate respectively a region of explosiveness from the region of indeterminacy, and another region of explosiveness from the region of determinacy.

It is important to notice that, in the presence of bounded rationality, the dynamics out of the determinacy region are quite different from the benchmark RE model. In fact, when boundedly rational agents are present in the economy, the model can show instability (i.e., explosive equilibria) instead of indeterminacy (i.e., multiple bounded equilibria).

As the fraction of rational agents decreases further, the determinacy region keeps on shrinking as shown in Figs. 3.2c and 3.2d. In fact, while Fig. 3.1a shows that the fold bifurcation curve approaches an asymptote, Fig. 3.1b shows that the flip bifurcation curve keeps on rotating counterclockwise as  $n_{RE}$  decreases, restricting the region of policy actions ensuring determinacy.

Notice also that the indeterminacy region shrinks as  $n_{RE}$  decreases. In fact Fig. 3.1c shows that, as the fraction of rational agents decreases, the Neimark-Sacker bifurcation surface associated with values of the policy reaction coefficient  $\phi_\pi < 1$  approaches the fold bifurcation surface of Fig. 3.1a.

Of course the results presented in Fig. 3.2 are not exhaustive of all possible determinacy scenarios, but they suffice to make clear the main point of our monetary policy exercise, namely that the presence of bounded rationality may alter significantly the determinacy properties of the model, and therefore that in a world with heterogeneous expectations an interest rate rule that obeys the Taylor principle does not necessarily guarantee a determinate equilibrium. As a concrete example, consider the interest rate rule with  $\phi_\pi = 1.5$  and  $\phi_y = 0.125$  proposed by

Taylor (1993)<sup>20</sup>. This policy rule ensures determinacy in the cases  $n_{RE} = 1$  and  $n_{RE} = 0.8$  (see top panels Fig. 3.2), however, when the fraction of rational agents decreases to  $n_{RE} = 0.6$  and  $n_{RE} = 0.4$ , the Taylor rule falls in the explosiveness region (see bottom panels Fig. 3.2). The intuition for such result can be found in the logic that produces unique stable dynamics in a New Keynesian model with homogeneous rational expectations. The stabilization mechanism after a shock in the New Keynesian model relies on unstable dynamics in the sense that, by obeying the Taylor principle, the monetary authority induces dynamics that will explode in any equilibrium but one. Ruling out explosive paths guarantees then uniqueness of output and inflation equilibrium paths<sup>21</sup>. However the presence of boundedly rational agents, whose expectations co-evolve with macroeconomic variables in a dynamic feedback system, introduces backward-looking components in the dynamics of the model. In the presence of backward-looking agents in the economy, parameters' regions that ensured unstable eigenvalues and thus determinate equilibrium in a completely forward-looking model, may now induce unstable dynamics. This result is of a crucial importance for the conduct of a sound monetary policy. In fact the rationale for recommendations advising to conduct policies within the determinacy region is based on the fact that determinacy reduces volatility of inflation and output. It is therefore very important to account for bounded rationality when designing monetary policy since policies constructed to achieve determinacy under homogeneous rational expectations may be destabilizing when expectations are heterogeneous.

### 3.5 Conclusions

Recent papers provided empirical evidence in favor of heterogeneity in individual expectations using survey data as well as experimental data. Building on this evidence, we derived a general micro-founded version of the New Keynesian framework for the analysis of monetary policy

---

<sup>20</sup>The original coefficient on the output term  $\phi_y = 0.5$  has been divided by 4 so that  $\phi_y$  corresponds to the output coefficient in a standard Taylor rule written in terms of annualized interest and inflation rates (see Woodford (2003)).

<sup>21</sup>See for example Cochrane (2010) for a discussion.

in the presence of heterogeneous expectations. We model individual behavior as being optimal given subjective expectations and derive a law of motion for output and inflation by explicitly aggregating individual decision rules. One advantage of our approach is that it is sufficiently general to consider a rich ecology of forecasting rules, ranging from simple heuristics to the very sophisticated rational expectation predictor. The RE benchmark is indeed a special case of our heterogeneous expectations New Keynesian model.

We designed an economy where agents can decide whether to pay some costs for rationality or use simple heuristics to forecast macroeconomic variables. After having characterized the dynamic feedback system in which aggregate variables and subjective expectations co-evolve over time, we performed a simple monetary policy exercise to illustrate the implications of expectations' heterogeneity on the determinacy properties of the model. Our central finding is that in a world with heterogeneous agents, the Taylor principle does not necessarily guarantee a unique equilibrium. Therefore policy makers should seriously take into account bounded rationality when designing monetary policies. In fact policy attempts to achieve determinacy under homogeneous rational expectations may destabilize the economy even when only a small fraction of boundedly rational agents is present in the economy.

This paper provides a theoretical DSGE framework for the analysis of monetary policy in the presence of heterogeneous beliefs. Future work should try to estimate the degree of rationality in the economy and investigate further the implications of heterogeneity in the way agents form their beliefs for the global dynamics of the economy and optimal monetary policy design.



## Appendix 3.A Steady state representative rational agent model

We consider a zero inflation steady state ( $\pi = 1$ ) around which we will linearize the equations of the model. The consumption Euler equation (3.3) implies then

$$R = \delta^{-1}.$$

Using the production function we get

$$y = h.$$

From the pricing equation (3.7) computed in steady state we get

$$w = \frac{\eta - 1}{\eta}.$$

Using that  $c = y = h$  in the labor supply function (3.4) implies

$$\chi h^{\gamma+\sigma} = \frac{\eta - 1}{\eta},$$

so normalizing  $\chi = \frac{\eta-1}{\eta}$  we have

$$h = c = y = 1.$$

Finally, using the definition of total dividends we get the steady state value

$$d = 1 - \frac{\eta - 1}{\eta} = \frac{1}{\eta}.$$

## Appendix 3.B Derivation of consumption rule (3.17)

Consider household  $i$ 's first order conditions

$$\hat{c}_{i,t} = \tilde{E}_{i,t} \hat{c}_{i,t+1} - \sigma^{-1} \left( \hat{R}_t - \tilde{E}_{i,t} \hat{\pi}_{t+1} \right) \quad (3.47)$$

$$\hat{h}_{i,t} = \gamma^{-1} (\hat{w}_t - \sigma \hat{c}_{i,t}) \quad (3.48)$$

and the intertemporal budget constraint

$$\tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \hat{c}_{i,s} = \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( \delta^{-1} \tilde{b}_{i,t-1} + (1 - \eta^{-1}) (\hat{w}_s + \hat{h}_{i,s}) + \eta^{-1} \hat{d}_s \right). \quad (3.49)$$

By substituting (3.48) into (3.49) we get

$$\tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \hat{c}_{i,s} = \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( \delta^{-1} \tilde{b}_{i,t-1} + (1 - \eta^{-1}) \hat{w}_s + \frac{(1 - \eta^{-1})}{\gamma} (\hat{w}_s - \sigma \hat{c}_{i,s}) + \eta^{-1} \hat{d}_s \right),$$

which can be rewritten as

$$\tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( 1 + \frac{(1 - \eta^{-1}) \sigma}{\gamma} \right) \hat{c}_{i,s} = \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( \delta^{-1} \tilde{b}_{i,t-1} + \frac{(1 - \eta^{-1}) (1 + \gamma)}{\gamma} \hat{w}_s + \eta^{-1} \hat{d}_s \right). \quad (3.50)$$

We can now iterate equation (3.47) to get

$$\tilde{E}_{i,t} \hat{c}_{i,s} = \hat{c}_{i,t} + \sigma^{-1} \tilde{E}_{i,t} \sum_{k=t}^s \left( \hat{R}_k - \hat{\pi}_{k+1} \right). \quad (3.51)$$

Substituting (3.51) in the LHS of (3.50) gives

$$\begin{aligned} & \left( 1 + \frac{(1 - \eta^{-1}) \sigma}{\gamma} \right) \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( \hat{c}_{i,t} + \sigma^{-1} \sum_{k=t}^s \left( \hat{R}_k - \hat{\pi}_{k+1} \right) \right) \\ &= \left( 1 + \frac{(1 - \eta^{-1}) \sigma}{\gamma} \right) \left( \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \hat{c}_{i,t} + \sigma^{-1} \tilde{E}_{i,t} \sum_{s=t}^{\infty} \sum_{k=t}^s \delta^{s-t} \left( \hat{R}_k - \hat{\pi}_{k+1} \right) \right) \\ &= \left( 1 + \frac{(1 - \eta^{-1}) \sigma}{\gamma} \right) \left( \frac{1}{1 - \delta} \hat{c}_{i,t} + \sigma^{-1} \tilde{E}_{i,t} \sum_{s=t}^{\infty} \sum_{k=t}^s \delta^{s-t} \left( \hat{R}_k - \hat{\pi}_{k+1} \right) \right). \quad (3.52) \end{aligned}$$

Since the term

$$\begin{aligned}
& \tilde{E}_{i,t} \sum_{s=t}^{\infty} \sum_{k=t}^s \delta^{s-t} \left( \hat{R}_k - \hat{\pi}_{k+1} \right) \\
&= \tilde{E}_{i,t} \left[ \begin{array}{c} \delta \left( \hat{R}_t - \hat{\pi}_{t+1} \right) + \\ \delta^2 \left( \hat{R}_t - \hat{\pi}_{t+1} \right) + \delta^2 \left( \hat{R}_{t+1} - \hat{\pi}_{t+2} \right) + \\ \delta^3 \left( \hat{R}_t - \hat{\pi}_{t+1} \right) + \delta^3 \left( \hat{R}_{t+1} - \hat{\pi}_{t+2} \right) + \delta^3 \left( \hat{R}_{t+2} - \hat{\pi}_{t+3} \right) + \\ \dots \end{array} \right] \\
&= \tilde{E}_{i,t} \left[ \begin{array}{c} (\delta + \delta^2 + \delta^3 + \dots) \left( \hat{R}_t - \hat{\pi}_{t+1} \right) + \\ (\delta^2 + \delta^3 + \delta^4 + \dots) \left( \hat{R}_{t+1} - \hat{\pi}_{t+2} \right) + \\ (\delta^3 + \delta^4 + \delta^5 + \dots) \left( \hat{R}_{t+2} - \hat{\pi}_{t+3} \right) + \\ \dots \end{array} \right] = \tilde{E}_{i,t} \left[ \begin{array}{c} \delta (1 + \delta + \dots) \left( \hat{R}_t - \hat{\pi}_{t+1} \right) + \\ \delta^2 (1 + \delta + \dots) \left( \hat{R}_{t+1} - \hat{\pi}_{t+2} \right) + \\ \delta^3 (1 + \delta + \dots) \left( \hat{R}_{t+2} - \hat{\pi}_{t+3} \right) + \\ \dots \end{array} \right] \\
&= \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta \frac{\delta^{s-t}}{1-\delta} \left( \hat{R}_s - \hat{\pi}_{s+1} \right),
\end{aligned}$$

we can rewrite (3.52) as

$$\left( 1 + \frac{(1-\eta^{-1})\sigma}{\gamma} \right) \left( \frac{1}{1-\delta} \hat{c}_{i,t} + \frac{\delta}{\sigma} \tilde{E}_{i,t} \sum_{s=t}^{\infty} \frac{\delta^{s-t}}{1-\delta} \left( \hat{R}_s - \hat{\pi}_{s+1} \right) \right). \quad (3.53)$$

Substituting now (3.53) into (3.50) we finally get a consumption rule for agent  $i$

$$\begin{aligned}
\hat{c}_{i,t} &= \frac{(1-\delta)\gamma}{\gamma + (1-\eta^{-1})\sigma} \left( \frac{1}{\delta(1-\delta)} \tilde{b}_{i,t-1} + \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \frac{(1-\eta^{-1})(1+\gamma)}{\gamma} \hat{w}_s + \eta^{-1} \hat{d}_s \right) \\
&\quad - \frac{\delta}{\sigma} \tilde{E}_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left( \hat{R}_s - \hat{\pi}_{s+1} \right). \quad (3.54)
\end{aligned}$$

### Appendix 3.C Derivation of pricing rule (3.22)

We can rewrite equation (3.19) as

$$p_{j,t}^{-\eta} \tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s \left( (1-\eta) \frac{p_{j,t}}{\pi_{t,s}} + \eta w_s \right) \pi_{t,s}^{\eta} c_s = 0$$

$\Leftrightarrow$

$$\tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s (\eta w_s \pi_{t,s}^{\eta} c_s) = \tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s ((\eta-1) \pi_{t,s}^{\eta-1} c_s p_{j,t})$$

$\Leftrightarrow$

$$p_{j,t} = \frac{\eta}{\eta-1} \frac{\tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s (\pi_{t,s}^{\eta} w_s c_s)}{\tilde{E}_{j,t} \sum_{s=t}^{\infty} \omega^{s-t} Q_s (\pi_{t,s}^{\eta-1} c_s)}$$

$\Leftrightarrow$

$$p_{j,t} \left( c_t + \tilde{E}_{j,t} \omega Q_{t+1} \pi_{t,t+1}^{\eta-1} c_{t+1} + \dots \right) = \frac{\eta}{\eta-1} \left( w_t c_t + \tilde{E}_{j,t} \omega Q_{t+1} \pi_{t,t+1}^{\eta} w_{t+1} c_{t+1} + \dots \right).$$

Log-linearizing gives

$$\begin{aligned} & (c + \omega \bar{Q}_{t+1} \pi^{\eta-1} c + \dots) (p_{j,t} - p) + p (c_t - c) + p \omega \bar{Q}_{t+1} \pi^{\eta-1} \tilde{E}_{j,t} (c_{t+1} - c) + \\ & p \omega \pi^{\eta-1} c \tilde{E}_{j,t} (Q_{t+1} - \bar{Q}_{t+1}) + (\eta-1) p \omega \pi^{\eta-2} c \tilde{E}_{j,t} (\pi_{t,t+1} - \pi) + \dots \\ = & \frac{\eta}{\eta-1} \left( c (w_t - w) + w (c_t - c) + \omega \bar{Q}_{t+1} \pi^{\eta} c \tilde{E}_{j,t} (w_{t+1} - w) + \omega \bar{Q}_{t+1} \pi^{\eta} w \tilde{E}_{j,t} (c_{t+1} - c) \right) + \\ & \omega \pi^{\eta} w c \tilde{E}_{j,t} (Q_{t+1} - \bar{Q}_{t+1}) + \eta \omega \bar{Q}_{t+1} \pi^{\eta-1} w c \tilde{E}_{j,t} (\pi_{t,t+1} - \pi) + \dots \end{aligned}$$

Using steady state values  $c = 1$ ,  $\pi = 1$ ,  $p = 1$ ,  $w = \frac{\eta-1}{\eta}$ ,  $\bar{Q}_s = \delta^{s-t}$  we can write

$$\begin{aligned} & (1 + \omega\delta + \omega^2\delta^2 + \dots) \hat{p}_{j,t} + \left( \hat{c}_t + \omega\delta \tilde{E}_{j,t} \hat{c}_{t+1} + \dots \right) + \left( \omega\delta \tilde{E}_{j,t} \hat{Q}_{t+1} + \dots \right) + \\ & (\eta - 1) \left( \omega\delta \tilde{E}_{j,t} \hat{\pi}_{t,t+1} + \omega^2\delta^2 \tilde{E}_{j,t} \hat{\pi}_{t,t+2} + \dots \right) \\ = & \left( \hat{w}_t + \omega\delta \tilde{E}_{j,t} \hat{w}_{t+1} + \dots \right) + \left( \hat{c}_t + \omega\delta \tilde{E}_{j,t} \hat{c}_{t+1} + \dots \right) + \left( \omega\delta \tilde{E}_{j,t} \hat{Q}_{t+1} + \dots \right) + \\ & \eta \left( \omega\delta \tilde{E}_{j,t} \hat{\pi}_{t,t+1} + \omega^2\delta^2 \tilde{E}_{j,t} \hat{\pi}_{t,t+2} + \dots \right) \end{aligned}$$

$\Leftrightarrow$

$$\frac{1}{1 - \omega\delta} \hat{p}_{j,t} - \left( \omega\delta \tilde{E}_{j,t} \hat{\pi}_{t,t+1} + \omega^2\delta^2 \tilde{E}_{j,t} \hat{\pi}_{t,t+2} + \dots \right) = \left( \hat{w}_t + \omega\delta \tilde{E}_{j,t} \hat{w}_{t+1} + \dots \right)$$

$\Leftrightarrow$

$$\frac{1}{1 - \omega\delta} \hat{p}_{j,t} = \tilde{E}_{j,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{w}_s + \omega\delta \tilde{E}_{j,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_{t,s+1}.$$

Using  $\pi_{t,s+1} = \pi_{s+1} \cdot \pi_s \cdot \dots \cdot \pi_{t+1}$  which means that  $\hat{\pi}_{t,s+1} = \hat{\pi}_{s+1} + \hat{\pi}_s + \dots + \hat{\pi}_{t+1}$  we have that

$$\begin{aligned} & \omega\delta \tilde{E}_{j,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_{t,s+1} \\ = & \omega\delta \tilde{E}_{j,t} \hat{\pi}_{t,t+1} + \omega^2\delta^2 \tilde{E}_{j,t} \hat{\pi}_{t,t+2} + \dots \\ = & \begin{bmatrix} \omega\delta \hat{\pi}_{t+1} + \\ \omega^2\delta^2 \hat{\pi}_{t+1} + \omega^2\delta^2 \hat{\pi}_{t+2} + \\ \omega^3\delta^3 \hat{\pi}_{t+1} + \omega^3\delta^3 \hat{\pi}_{t+2} + \omega^3\delta^3 \hat{\pi}_{t+3} + \\ \dots \end{bmatrix} = \begin{bmatrix} \omega\delta (1 + \omega\delta + \omega^2\delta^2 + \dots) \hat{\pi}_{t+1} + \\ \omega^2\delta^2 (1 + \omega\delta + \omega^2\delta^2 + \dots) \hat{\pi}_{t+2} + \\ \omega^3\delta^3 (1 + \omega\delta + \omega^2\delta^2 + \dots) \hat{\pi}_{t+3} + \\ \dots \end{bmatrix} \\ = & \frac{\omega\delta}{1 - \omega\delta} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_s. \end{aligned}$$

We can thus write

$$\frac{1}{1 - \omega\delta} \hat{p}_{j,t} = \tilde{E}_{j,t} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{w}_s + \frac{\omega\delta}{1 - \omega\delta} \sum_{s=t}^{\infty} (\omega\delta)^{s-t} \hat{\pi}_s,$$

which leads to equation (3.22).

### Appendix 3.D Derivation of systems (3.43) and (3.46)

Start from system (3.40) which can be rewritten using results (3.41) and (3.42) as

$$\begin{aligned}\widehat{y}_t &= n_{RE,t} E_t \sum_{s=t}^{\infty} \delta^{s-t} \left( (1-\delta)\widehat{y}_s - \frac{\delta}{\sigma}(\widehat{R}_s - \widehat{\pi}_{s+1}) \right) + \\ &\quad (1 - n_{RE,t}) \left( \widehat{y}_{t-1} - \frac{\delta}{\sigma}\widehat{R}_t - \frac{\delta^2}{(1-\delta)\sigma}\widehat{R}_{t-1} + \frac{\delta}{(1-\delta)\sigma}\widehat{\pi}_{t-1} \right) \\ \widehat{\pi}_t &= n_{RE,t} E_t \sum_{s=t}^{\infty} (\omega\delta)^{s-t} (k\widehat{y}_s + (1-\omega)\delta\widehat{\pi}_{s+1}) + (1 - n_{RE,t}) \left( \frac{k}{1-\omega\delta}\widehat{y}_{t-1} + \frac{(1-\omega)\delta}{1-\omega\delta}\widehat{\pi}_{t-1} \right).\end{aligned}\tag{3.55}$$

Therefore we have that  $V_x(\cdot)$ , for  $x \in \{y, \pi\}$ , in (3.43) are defined as

$$V_y(\widehat{y}_{t-1}, \widehat{R}_t, \widehat{R}_{t-1}, \widehat{\pi}_{t-1}) = \widehat{y}_{t-1} - \frac{\delta}{\sigma}\widehat{R}_t - \frac{\delta^2}{(1-\delta)\sigma}\widehat{R}_{t-1} + \frac{\delta}{(1-\delta)\sigma}\widehat{\pi}_{t-1} \tag{3.56}$$

$$V_\pi(\widehat{y}_{t-1}, \widehat{\pi}_{t-1}) = \frac{k}{1-\omega\delta}\widehat{y}_{t-1} + \frac{(1-\omega)\delta}{1-\omega\delta}\widehat{\pi}_{t-1} \tag{3.57}$$

As standard in the learning literature we assumed that the current interest rate is observed by non-rational agents while current output and inflation are not. The term  $\widehat{R}_{t-1}$  shows up in the aggregate demand via the performance measure (3.34) in the heuristics selection process. This is due to the assumption that the selection of heuristics takes place at the beginning of period  $t$ , before observing  $\widehat{R}_t$ .

Using the result (3.45) to drop the time index from the term  $n_{RE,t}$  and closing the model with the interest rate rule  $\widehat{R}_t = \phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t$ , after some algebraic manipulations we can rewrite the system (3.55) as

$$E_t \widehat{y}_{t+1} = q_{yy} \widehat{y}_t + q_{yE\pi} E_t \widehat{\pi}_{t+1} + q_{y\pi} \widehat{\pi}_t + q_{yY1} \widehat{y}_{t-1} + q_{y\pi1} \widehat{\pi}_{t-1} \tag{3.58}$$

$$E_t \widehat{\pi}_{t+1} = \gamma_{\pi y} \widehat{y}_t + \gamma_{\pi\pi} \widehat{\pi}_t + \gamma_{\pi Y1} \widehat{y}_{t-1} + \gamma_{\pi\pi1} \widehat{\pi}_{t-1} \tag{3.59}$$

with

$$\begin{aligned}
\Omega_y &= \left(1 + \frac{\delta(1 - n_{RE})}{\sigma} \phi_y\right)^{-1} \delta^{-1} \\
q_{yy} &= \Omega_y \left(1 - n_{RE} + \delta - \left(\frac{\delta^3(1 - n_{RE})}{(1 - \delta)\sigma} - \frac{\delta}{\sigma}\right) \phi_y\right) \\
q_{yE\pi} &= \Omega_y \left(-\frac{n_{RE}\delta}{\sigma} - \frac{\delta^2(1 - n_{RE})}{\sigma} \phi_\pi\right) \\
q_{y\pi} &= \Omega_y \left(\frac{\delta^2(1 - n_{RE})}{(1 - \delta)\sigma} - \left(\frac{\delta^3(1 - n_{RE})}{(1 - \delta)\sigma} - \frac{\delta}{\sigma}\right) \phi_\pi\right) \\
q_{yY1} &= \Omega_y \left(-(1 - n_{RE}) \left(1 - \frac{\delta^2}{(1 - \delta)\sigma} \phi_y\right)\right) \\
q_{y\Pi1} &= \Omega_y \left(-(1 - n_{RE}) \left(\frac{\delta}{(1 - \delta)\sigma} - \frac{\delta^2}{(1 - \delta)\sigma} \phi_\pi\right)\right) \\
\Omega_\pi &= (n_{RE}(1 - \omega)\delta + \omega\delta)^{-1} \\
\gamma_{\pi y} &= \Omega_\pi \left(-n_{RE}k + \frac{\omega\delta(1 - n_{RE})}{1 - \omega\delta}\right) \\
\gamma_{\pi\pi} &= \Omega_\pi \left(1 + \frac{(1 - \omega)\omega\delta^2(1 - n_{RE})}{1 - \omega\delta}\right) \\
\gamma_{\pi Y1} &= \Omega_\pi \left(-(1 - n_{RE}) \left(\frac{k}{1 - \omega\delta}\right)\right) \\
\gamma_{\pi\Pi1} &= \Omega_\pi \left(-(1 - n_{RE}) \frac{(1 - \omega)\delta}{1 - \omega\delta}\right)
\end{aligned}$$

Plugging (3.59) into (3.58) and rearranging terms we finally get

$$\begin{aligned}
E_t \hat{y}_{t+1} &= \gamma_{yy} \hat{y}_t + \gamma_{y\pi} \hat{\pi}_t + \gamma_{yY1} \hat{y}_{t-1} + \gamma_{y\Pi1} \hat{\pi}_{t-1} \\
E_t \hat{y}_{t+1} &= \gamma_{\pi y} \hat{y}_t + \gamma_{\pi\pi} \hat{\pi}_t + \gamma_{\pi Y1} \hat{y}_{t-1} + \gamma_{\pi\Pi1} \hat{\pi}_{t-1}
\end{aligned} \tag{3.60}$$

where

$$\begin{aligned}
\gamma_{yy} &= q_{yy} + q_{yE\pi} \gamma_{\pi y} \\
\gamma_{y\pi} &= q_{y\pi} + q_{yE\pi} \gamma_{\pi\pi} \\
\gamma_{yY1} &= q_{yY1} + q_{yE\pi} \gamma_{\pi Y1} \\
\gamma_{y\Pi1} &= q_{y\Pi1} + q_{yE\pi} \gamma_{\pi\Pi1}.
\end{aligned}$$

System (3.60) can be rewritten in matrix form as in (3.46).

## Appendix 3.E Computation of bifurcation surfaces

The characteristic equation of matrix  $\mathbb{M}$  is given by

$$\mathcal{P}(\lambda) = \lambda^4 + c_3\lambda^3 + c_2\lambda^2 + c_1\lambda + c_0 = 0 \quad (3.61)$$

where

$$c_3 = -\gamma_{yy} - \gamma_{\pi\pi}$$

$$c_2 = \gamma_{yy}\gamma_{\pi\pi} - \gamma_{y\pi}\gamma_{\pi y} - \gamma_{yY1} - \gamma_{\pi\Pi1}$$

$$c_1 = -\gamma_{y\pi}\gamma_{\pi Y1} - \gamma_{\pi y}\gamma_{y\Pi1} + \gamma_{yy}\gamma_{\pi\Pi1} + \gamma_{\pi\pi}\gamma_{yY1}$$

$$c_0 = \gamma_{yY1}\gamma_{\pi\Pi1} - \gamma_{y\Pi1}\gamma_{\pi Y1}.$$

The computation of the fold and flip bifurcation surfaces are pretty straightforward. By plugging  $\lambda = 1$  and  $\lambda = -1$  in (3.61), one finds that the fold and the flip bifurcations loci are given respectively by

$$1 + c_3 + c_2 + c_1 + c_0 = f_{\text{fold}}(\phi_\pi, \phi_y, n_{RE}) = 0$$

$$1 - c_3 + c_2 - c_1 + c_0 = f_{\text{flip}}(\phi_\pi, \phi_y, n_{RE}) = 0.$$

The locus of Neimark-Sacker bifurcations is associated with complex roots of modulus 1,  $\lambda_{1,2} = e^{\pm i\theta}$ , where  $\theta$  is an arbitrary angle and  $i$  is the imaginary unit. Using Vieta's formulas we find the following relations between the coefficients of  $\mathcal{P}(\lambda)$  and the roots of (3.61):

$$c_3 = -\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 \quad (3.62)$$

$$c_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4 \quad (3.63)$$

$$c_1 = -\lambda_1\lambda_2\lambda_3 - \lambda_2\lambda_3\lambda_4 - \lambda_1\lambda_3\lambda_4 - \lambda_1\lambda_2\lambda_4 \quad (3.64)$$

$$c_0 = \lambda_1\lambda_2\lambda_3\lambda_4. \quad (3.65)$$



Given that  $\lambda_1\lambda_2 = e^{i\vartheta}e^{-i\vartheta} = 1$ , we can rewrite (3.65) as

$$c_0 = \lambda_3\lambda_4. \quad (3.66)$$

Moreover, using (3.62) we have that the sum of the real roots is given by

$$\lambda_3 + \lambda_4 = -c_3 - e^{i\vartheta} - e^{-i\vartheta}. \quad (3.67)$$

By plugging (3.66) and (3.67) in (3.64) and (3.63) we get

$$c_1 = -c_0(e^{i\vartheta} + e^{-i\vartheta}) + c_3 + e^{i\vartheta} + e^{-i\vartheta} \quad (3.68)$$

$$c_2 = 1 + c_0 + e^{i\vartheta}(-c_3 - e^{i\vartheta} - e^{-i\vartheta}) + e^{-i\vartheta}(-c_3 - e^{i\vartheta} - e^{-i\vartheta}). \quad (3.69)$$

Using Euler's formula we can rewrite (3.68) and (3.69) as

$$c_1 = (1 - c_0)2 \cos(\vartheta) + c_3 \quad (3.70)$$

$$c_2 = 1 + c_0 + 2 \cos(\vartheta)(-c_3 - 2 \cos(\vartheta)). \quad (3.71)$$

Plugging (3.70) in (3.71) we have that the locus of Neimark-Sacker bifurcation is given by

$$1 + c_0 - c_2 - \frac{c_3(c_1 - c_3)}{1 - c_0} - \left( \frac{c_1 - c_3}{1 - c_0} \right)^2 = f_{NS}(\phi_\pi, \phi_y, n_{RE}) = 0$$

subject to  $-1 \leq (c_1 - c_3)/(2(1 - c_0)) \leq 1$ .