Tilburg University

Formation of Rationally Heterogeneous Expectations

Pfajfar, D.

Publication date:
2012

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Pfajfar, D. (2012). *Formation of Rationally Heterogeneous Expectations*. (CentER Discussion Paper; Vol. 2012-083). Economics.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2012-083

**FORMATION OF RATIONALLY
HETEROGENEOUS EXPECTATIONS**

By

Damjan Pfajfar

22 October, 2012

ISSN 0924-7815

Formation of Rationally Heterogeneous Expectations*

Damjan Pfajfar[†]

CentER, EBC, University of Tilburg

October 14, 2012

Abstract

This paper models expectation formation by taking into account that agents produce heterogeneous expectations due to model uncertainty, informational frictions and different capacities for processing information. We show that there are two general classes of steady states within this framework: those where strictly dominated forecasting rules vanish, and intrinsic heterogeneous steady states where a positive proportion of agents use a more costly perfect foresight. This demonstrates that intrinsic heterogeneity can also arise in a model where the forecasting rules are not equally costly, and do not exhibit identical performance in the long run.

JEL Classification: C62; D83; D84; E30

Keywords: Heterogeneous expectations, cobweb model, adaptive learning, rational expectations.

*I would like to thank two anonymous referees, the editor, Seppo Honkapohja, Cars Hommes, Chryssi Giannitsarou, Johannes Binswanger and other seminar participants at Cambridge University, REPLHA International Conference in Milan and Learning week at St. Louis FED for helpful comments. Any remaining errors are my own.

[†]Department of Economics, Faculty of Economics and Business Administration, P.O. Box 90153, NL-5000 LE Tilburg, Netherlands. *E-mail:* D.Pfajfar@uvt.nl. *Tel.:* +31 13 466 3420. *Fax:* +31 13 466 3042. *Web:* <https://sites.google.com/site/dpfajfar/>.

1 Introduction

New evidence on the limitations of the rational expectations hypothesis has lead researchers to re-examine the process of expectation formation and its implications for macroeconomic models. In the theoretical literature, Evans and Honkapohja (2001) suggest the adaptive learning approach, i.e., agents act as econometricians when forecasting. In these models agents know the correct model of the economy but do not know the parameters. In addition, bounded rationality has also been imposed in the form of sticky information models, e.g., Ball et al. (2005), where agents update information infrequently, and consequently their forecasts produce systematic errors. However, different papers have advanced evidence of the heterogeneity of expectations recently (for a survey see Hommes, 2011). Generally, three main sources of heterogeneity have been documented: (i) agents produce heterogeneous forecasts because they are using different models, (ii) they have different information sets and (iii) they have different capacities for processing information. Furthermore, there are different reasons for agents to use different models. They can either believe in different structure of the economy or they are "forced" to resort to simpler models because of either (ii) or (iii). The latter represents the focus of our paper. Theoretical models where agents are allowed to choose from a set of alternative models to forecast are usually labelled as rationally heterogeneous expectations models. They explore whether it is always optimal, from a utility maximization standpoint, to follow a rational (or more generally sophisticated) way of forming beliefs as they entail a fixed cost for processing information. Typically, in these frameworks the choice of belief formation is based on predictors' past performance. In this scenario there might be a situation in which a fraction of the population prefers to commit systematic forecast errors. Examples of these models are Brock and Hommes (1997), Branch and McGough (2008) and Waters (2009).

In this paper we develop a model of a dynamic predictor that takes into account several roots of heterogeneity. Our goal is to show that agents use different models for forecasting in equilibrium because they have different capacities to process information and different information sets. Branch and Evans (2006) show that intrinsic heterogeneity is possible among agents with similar skills and abilities who choose predictors (that are costless) from a set of optimal misspecified econometric models, which produce identical long run performance. We demonstrate that intrinsic heterogeneity can arise even among agents with different skills who use alternative rules that do not perform equally well in the long run, and there is a cost attached

to the more sophisticated rule. Heterogeneity can also emerge in the Brock and Hommes (1997) setup under finite intensity of choice parameter, which is consistent with boundedly rational behavior. It is more difficult to present a model where heterogeneity arises also under an infinitely high intensity of choice parameter as in Branch and Evans (2006 and 2007). In this case agents rationally decide which model to use, and simple balancing of costs and benefits will not do the job in a model with self referential properties as the cobweb model. While the Branch and Evans (2006 and 2007) approach shows that heterogeneity might arise between agents that have similar economic skills (but believing in different models), we show that heterogeneity might also emerge between agents with different abilities or different economics skills (literacy), under both finite and infinite intensities of choice parameter. In our framework, not all agents have the capacity to perform the computations needed to estimate, for example, a vector autoregression model. Maintaining generality, we can also observe convergence to a strictly dominating predictor, depending on the parameterization of the model. To compare our results with the existing literature on the intrinsic heterogeneity, we investigate a special case of the choice between perfect foresight and naive expectations in a stochastic version of the cobweb model.

Compared to Brock and Hommes (1997), Branch and McGough (2008) and Waters (2009), we present a dynamic predictor selection that takes into account different roots of potential heterogeneity. Only a limited number of individuals can use the rational predictor as not all of them have the sufficient capacity to process information to form rational beliefs (perfect foresight). We believe that investing in knowledge or skills (capacity to process information) is a long term decision. Other predictors commonly used in economics fail to take into account that some agents might not be able to switch immediately to a sophisticated predictor. In addition to long run decisions regarding the capacity to process information, we also model information updating, which is a short horizon decision. Agents have to regularly update their information sets to be able to forecast with perfect foresight. The predictor put forward in this paper is determined in two stages. In the first stage, agents choose between the costly building up of the capacity to process information and free riding, which decreases their capacity to process information. This stage is modeled by adopting one of the standard predictors in the literature. In the basic example we use New Learning Dynamics (NLD), as in Waters (2009), but the generalized replicator dynamics, as in Branch and McGough (2008), could be used as well without changing the main results. In the second stage, we model information acquisition, though in a very simple way. Even when agents build up capacity they might not observe

all the relevant information to use the perfect foresight.¹ The latter root of heterogeneity has received a lot of attention in macroeconomics as it can help to reconcile standard models with the stylized facts, see e.g., Ball et al. (2005) and Branch et al. (2009).

The idea of ability constraints, and availability of information, was discussed in the economics literature of the late 19th and early 20th centuries and in the economics psychology literature at least from Herbert Simon onwards (Whalley, 2005). Heckman et al. (2006) shows that cognitive and noncognitive skills explain several labor market decisions as well as schooling decisions. A potential shortcoming of previous models of dynamic predictor selection is their assumption that all agents are capable of using the perfect foresight at any point of time. However, in reality they have to learn and build up capacity to process information for a long period of time before they can use the perfect foresight. It is possible to think about that as investing in human capital, as e.g., investing into education or more relevant to our case economic literacy. Although, by construction, all economic agents in our model have similar cognitive abilities as all of them can eventually build up the capacity to a level where they can use a sophisticated predictor. Burke and Manz (2011) show that subjects with higher economic literacy both better choose the information to use for forecasting and better use given information in the inflation forecasting experiment. Some empirical support that education is an important determinant of the accuracy of inflation expectations is put forward by Bryan and Venkatu (2001), Souleles (2004), Pfajfar and Santoro (2009) and Anderson et al. (2010) for the Michigan survey of Consumers, where on average 500 telephone owning households in the US are interviewed every month. These studies assess inflation expectations across different demographic groups and find significant differences among them. These demographic variables could serve as a proxy for the capacity to process information (see e.g., Heckman et al., 2006, Dohmen et al., 2010, and Grinblatt et al., 2011). Highly educated agents (and those with higher income) produce significantly lower forecast errors and they update their information sets more regularly. Also there is evidence using the same dataset that agents do not frequently update their information sets (see e.g., Mankiw et al., 2004, Carroll, 2003, Coibion and Gorodnichenko, 2012). Therefore, informational constraints are also binding for a significant proportion of households.

Different models of prediction selection are used in economics.² Brock and Hommes (1997) model predictor selection using a multinomial rule in a deterministic cobweb model and find

¹Figure A1 in the Appendix provides a simple explanation of the decision process.

²These models originate from population ecology and population genetics models, see Schuster and Sigmund (1983), and have been first implemented in economics in game theoretical studies.

heterogeneity as an equilibrium outcome. For a deterministic cobweb model with the choice between the perfect foresight and naive expectations they observe that when all agents use the cheapest predictor, the steady state is unstable. In the case of a strictly dominating rational (costly) predictor, the steady-state is stable. The asynchronous updating version of the Brock and Hommes (1997) predictor is proposed in Hommes et al. (2005a), where only a constant fraction of agents can update their choice of predictor every period. In contrast, Branch and McGough (2008) model prediction selection with a version of replicator dynamics which has fixed points on the boundaries of the simplex. As in Brock and Hommes (1997), they find evidence that when all agents use the cheapest predictor the steady state is unstable. Also the New Learning Dynamic of Brown and von Neumann (1950) has been discussed recently within a macro framework. In this setup, as it is outlined in Waters (2009), the best performing strategy has a monotonically increasing fraction of followers.

As Branch and McGough (2008) point out, one possible criticism of the Brock and Hommes (1997) approach is that they obtain a steady state where there exists a positive proportion of agents that use a strictly dominated predictor.³ However, in reality it is clear that not all agents can use the perfect foresight; they do not have sufficient economic knowledge to do that. Evidence from household inflation expectations surveys suggest that inflation expectations are on average biased (see Soules, 2004 and Pfajfar and Santoro, 2010 for evidence on Michigan survey data).⁴ Branch (2004, 2007) assesses the importance of informational frictions and model uncertainty using the Michigan survey of consumers, and finds that data are consistent with both of them. These models are capable of replicating some characteristics of the empirical distribution of forecasts. Pfajfar and Santoro (2010) analyze the same dataset on inflation expectations. They identify three regions of the overall distribution corresponding to different groups of expectation formation processes: a static or highly autoregressive group on the right-hand side of the median of the distribution, a nearly rational group around the median, and a group of agents on the right-hand side of the median that behave in accordance with adaptive learning and sticky information. The relationship between observing news and inflation expectation is detailed in Pfajfar and Santoro (2012), where it is shown that in the Michigan survey observing news on inflation increases the probability of updating inflation expectations.

³In their model only in the case of infinite intensity of choice parameter does the share of users converge to the strictly dominating predictor.

⁴The evidence exists also for surveys of consumers in Sweden (for both, Swedish Consumer Tendency Survey and Households Purchasing Plans Survey), e.g., for this evidence on the bias of inflation expectations of consumers see Jonung (1981), Bryan and Palmqvist (2005), and Dräger (2011).

Thus, this paper does not argue that the interior steady state is undesirable, *per se*. However, the predictor has to have the property that in periods when inflation matters – i.e., in periods of high inflation volatility – increasing number of agents use the more accurate predictor. Furthermore, it should drive out of the market the strictly dominated predictors that are not competitive (in the long run) with respect to other predictors. This property has been labelled as positive correlation in Sandholm (2010) and Waters (2009). We design our dynamic predictor in order to exhibit these three characteristics, as they are confirmed by the various empirical studies mentioned above.

Heterogeneous expectations and switching based on a multinomial logit have been recently introduced into the standard New Keynesian macro model by Branch and McGough (2009). Branch (2007) provides empirical evidence supporting the basic mechanism of the rationally heterogeneous expectations models within New Keynesian models. Pfajfar and Žakelj (2012), Assenza et al. (2011) and Anufriev and Hommes (2009) further investigate potential switching between different rules in an experimental framework. While Pfajfar and Žakelj (2012) focus on switching without a particular structure, Assenza et al. (2011) and Anufriev and Hommes (2009) employ a version of the predictor dynamics analyzed in Hommes et al. (2005a). They show that in an asset pricing environment the model with switching between simple heuristic rules can replicate the main results of the Hommes et al. (2005b) experiment in terms of individual behavior and aggregate dynamics. Pfajfar and Žakelj (2012) find that on average agents switch between different models every four periods. There is a strong evidence of intrinsic heterogeneity in groups consisting of nine subjects, as on average 4.5 different models are used in every period. Therefore, in both survey and experimental data heterogeneity is pervasive.⁵

The resulting dynamic predictor selection combines the properties of Adaptively Rational Equilibrium Dynamics (ARED) and Replicator Equilibrium Dynamics (RED). It produces a steady state where all agents converge towards a strictly dominating predictor (as in the RED) as well as steady states where both rules coexist (ARED). Another important result in our paper is that we can obtain intrinsic heterogeneity without restricting our choice to different costless or equally costly choices, as in Branch and Evans (2006). Our results do not depend on the intensity of choice parameter, as in Brock and Hommes (1997), thus we can observe intrinsic heterogeneity for infinite values of the intensity of choice parameter. We analyze these results analytically and provide intuition. Results also suggest that when all agents choose the

⁵See Hommes (2011) for a survey on experimental evidence of heterogeneous expectations.

simplest predictor, we can confirm the results of Brock and Hommes (1997) and Branch and McGough (2008) that this steady state is unstable for a high enough ratio between demand and supply slopes. Under this assumption we observe chaotic behavior. However, due to the complexity of our approach it is difficult to analyze the dynamics of the model using analytical tools. The main problem is that our system might not be differentiable over the whole of the relevant space. Furthermore, it is possible to show that other predictors emerge as a special case of the model presented in this paper.

The paper is organized as follows: Sections 2 and 3 discuss a general version of the cobweb model and the predictor with two different forecasting rules. Steady states and the dynamics of the model are then analyzed in Section 4. Section 5 concludes.

2 Cobweb Model with Perfect Foresight vs. Naive Expectations

In the setup below we focus on a specific set of choices, where agents choose between the perfect foresight and the naive or myopic predictor.⁶ The model discussed here is a version of the stochastic cobweb model, which consists of demand and supply equations:

$$D(p_t) = m_I - Bp_t, \quad (1)$$

$$S(p_{t+1}^e) = r_I + bH_h(p^t) + v_t, \quad (2)$$

where p_t is price at time t and p_{t+1}^e are price expectations at time t for $t+1$. $B, b > 0$, m_I and r_I are intercepts and v_t is white noise. Without loss of generality to the stability properties we can set m_I and r_I to 0. $\alpha = b/B$ is the ratio of the supply and demand slopes, which governs the self referential property of the model. We assume quadratic costs in the general form $\frac{1}{2}G(Q_t^*)^2$, where $G > 0$ and the total quantity is equal to the planned quantity $Q_t = Q_t^* + \kappa_t$ adjusted for exogenous productivity shocks realized after the production decisions are made. The corresponding problem of profit maximization for this setting takes the form:

$$\arg \max_{Q_t^*} \pi_t = E_t \left[p_{t+1}(Q_t) - \frac{1}{2}G(Q_t^*)^2 \right],$$

where κ_t is iid with zero mean. The solution to this problem leads to a supply equation $Q_t = \frac{1}{G}p_{t+1}^e + \kappa_t$, where p_{t+1}^e is equal to $E_t p_{t+1}$. Moreover, if we compare this solution to

⁶In a working paper version available on request from the authors, we present a generalized predictor with an arbitrary number of different rules.

equation (2), we get the following form of quadratic costs:

$$C_t = \frac{(Q_t^*)^2}{2b}.$$

2.1 Expectation Formation

To close the model we need to specify the processes of expectation formation. As in Brock and Hommes (1997) and in Branch and McGough (2008) we assume that agents have the choice between the perfect foresight and the naive predictor:

$$H_1(p^t) = p_{t+1}, \quad (3)$$

$$H_2(p^t) = p_t. \quad (4)$$

Proportion of agents n_t use the $H_1(p^t)$, while the rest of agents $(1 - n_t)$ use $H_2(p^t)$.

2.2 Heterogeneous Expectations Equilibrium

Applying market clearing condition, $D(p_{t+1}) = S(p_{t+1}^e)$, and taking into account the expectation formation mechanisms we get the following actual law of motion (ALM):

$$p_{t+1} = \frac{(n_t - 1)bp_t}{B + n_tb} - \frac{1}{B + n_tb}v_t. \quad (5)$$

Heterogeneous Expectations Equilibrium (HEE) is defined as a process (5), where $n_t = n$ is fixed. For a given n an HEE exists and is unique.

Assumption 1 *We consider the case where $B, b > 0$, and $\alpha (= b/B) > 0$.*

Note that Assumption 1 is also a necessary and sufficient condition that a process (5) represents a well defined mapping for all n in the unit simplex $S = (n \in \mathbb{R} : 0 \leq n \leq 1)$. Assumption 1 also restricts this analysis to the case when the cobweb model exhibits a negative feedback. Note that the process in (5) is a (non-homogeneous and non-autonomous) first order difference equation, thus it has a solution (for a given n):

$$p_{t+1} = \begin{cases} \left(\frac{(n-1)b}{B+nb}\right)^{t+1} p_0 - \sum_{s=0}^{t+1} \left(\frac{(n-1)b}{(B+nb)^2}\right)^{t+1-s} v_s & \text{if } \left|\frac{(n-1)b}{(B+nb)^2}\right| \leq 1 \\ \left(\frac{(n-1)b}{B+nb}\right)^{t+1} p_0 - \sum_{s=t+1}^{\infty} \left(\frac{1}{(n-1)b}\right)^{s-t-1} v_s & \text{if } \textit{otherwise} \end{cases}. \quad (6)$$

We can immediately observe that in the case when $\left|\frac{(n-1)b}{(B+nb)^2}\right| > 1$ instability of the equilibrium prices can emerge as previously pointed out by Brock and Hommes (1997) and Branch and

McGough (2008). Next our focus is shifted to the determination of the n , so that we can analyze the properties of the model and equilibria where n_t is determined endogenously. The latter equilibrium is called a Misspecification Heterogeneous Expectations Equilibrium.

3 Predictor Choice Dynamics

This paper introduces a predictor choice dynamics in line with Brock and Hommes (1997), Branch and McGough (2008), and Waters (2009), and augments it for the capacity to process information and limited distribution of information in the sense of Ball et al. (2005). Thus, it follows the recent literature on inattention. Nevertheless, contrary to previous literature, this paper divides the whole process into two separate stages. The first one consists of processing the observed information for which agents require a critical mass of knowledge in order to be able to forecast with the perfect foresight. In each period agents are assumed to make a choice between either investing a small cost in building up the capacity, or free riding. Agents in the rational group can free ride only for a limited number of periods, while it is assumed that they lose some small amount of knowledge (skills) every period, they do not choose to build up the capacity. This could be justified, e.g., if agents do not observe the technological progress or new methods developed in econometrics. Perfect foresight is always the dominant strategy if agents have a high enough capacity to process information. The second stage in the process consists of observing the information. For simplicity we assume that only a constant fraction of agents observe all the necessary information to be able to use perfect foresight. However, it can be straightforwardly endogenized in line with a paper by Branch et al. (2009).⁷

Comparing our approach with papers by Ball et al. (2005) and Mankiw and Reis (2001), this paper argues that the process of observing and the actual processing of information can be split into two separate stages. It might not be obvious why we endogenize the process of observing the information and the stage of deciding to update the capacity to process information as separate stages. It would not be efficient to update the capacity to process information (and pay a fixed cost to do that) while not paying the fixed cost for observing information. However, it would make sense if some agents, whose capacity to process information is relatively low, decide that they would like to increase their capacity in order to have the opportunity in the future to use the more efficient predictor. In the early stages of this process it may be efficient to

⁷Alternatively it could be endogeneized as well by employing any predictor selection process. However, in that case the appropriate fitness function is current profits and not average profits.

only pay the cost for building up the capacity and not the costs for observing the information. In other words, the two processes are inherently different, as the decision about the capacity to process information is a long term decision, while updating the information set is a short term one. This explains why it is useful to model these two processes separately, contrary to the existing literature on inattentiveness.

	Rational group		Boundedly rational group
	Perfect foresight	Naive rule	Naive rule
Building up capacity	$n_t \eta_{1,t}^a$	$\left(\frac{1-\lambda}{\lambda}\right) n_t \eta_{1,t}^a$	$\left(\frac{\lambda-n_t}{\lambda}\right) \eta_{2,t}^a$
Free riding	$n_t(1 - \eta_{1,t}^a)$	$\left(\frac{1-\lambda}{\lambda}\right) n_t(1 - \eta_{1,t}^a)$	$\left(\frac{\lambda-n_t}{\lambda}\right)(1 - \eta_{2,t}^a)$

Table 1: Groups of agents and predictors used

There are two groups of agents: the rational group and the boundedly rational group. Parameter λ denotes the proportion of agents who receive all the necessary information in order to be able to forecast with perfect foresight. Thus, the actual proportion of agents that are in the rational group is $\frac{n_t}{\lambda}$. We have two subgroups of agents in this group who either choose to build up the capacity or to free ride in each period, but still have the capacity to process information at a sufficiently high level. Agents who do not observe all the relevant information, but have high enough capacity to process information to access the perfect foresight still stay in the rational group but they cannot use the perfect foresight. Instead they have to resort to the naive predictor. Therefore agents in the rational group use both the perfect foresight and the naive predictor. In the boundedly rational group all agents use the naive predictor. Also in the boundedly rational group there are two subgroups. Agents in this group either choose to build up the capacity, and have the capacity to process information below the threshold level, or to free ride. η_h^a is an "adjusted" proportion of agents (for entry and exit, see below) that build up the capacity to process information in group h , where $h = 1$ corresponds to the rational group and $h = 2$ to the boundedly rational group. Note that η_h is a proportion of agents that build up the capacity to process information in group h before the entry and exit takes place. We can start by defining the law of motion for the proportion of agents that use perfect foresight, n_t :

$$\begin{aligned}
n_{t+1} &= \lambda(\Gamma_{1,t+1} + \Gamma_{2,t+1}) = n_t + \lambda(\Xi_{3,t+1} - \Gamma_{3,t+1}), \\
\Gamma_{1,t+1} &= \eta_{1,t+1} \left(\frac{n_t}{\lambda}\right) + \Xi_{3,t+1} = \eta_{1,t+1}^a \left(\frac{n_{t+1}}{\lambda}\right), \\
\Gamma_{2,t+1} &= (1 - \eta_{1,t+1}) \left(\frac{n_t}{\lambda}\right) - \Gamma_{3,t+1} = (1 - \eta_{1,t+1}^a) \left(\frac{n_{t+1}}{\lambda}\right).
\end{aligned} \tag{7}$$

The rational group is comprised by agents who build up the capacity and those who free

ride. $\Gamma_{1,t+1}(\Gamma_{2,t+1})$ is the proportion of agents that build up capacity (free ride) and are in the rational group in the period $t + 1$. $\Xi_{3,t+1}$ is the proportion of agents that have been upgraded from the boundedly rational group to the rational group in the period $t + 1$, as their capacity to process information reaches the critical level. $\Gamma_{3,t+1}$ is the proportion of agents that are relegated from the rational group in the period $t + 1$ as their capacity to process information falls under the critical level. The boundedly rational group is defined as follows:

$$\begin{aligned}\Xi_{1,t+1} &= (1 - \eta_{2,t+1}) \left(1 - \frac{n_t}{\lambda}\right) + \Gamma_{3,t+1} = (1 - \eta_{2,t+1}^a) \left(1 - \frac{n_{t+1}}{\lambda}\right), \\ \Xi_{2,t+1} &= \eta_{2,t+1} \left(1 - \frac{n_t}{\lambda}\right) - \Xi_{3,t+1} = \eta_{2,t+1}^a \left(1 - \frac{n_{t+1}}{\lambda}\right).\end{aligned}\tag{8}$$

There are two subgroups of boundedly rational group. $\Xi_{2,t+1}(\Xi_{1,t+1})$ is the proportion of agents that build up the capacity (free ride) and are in the bounded rational group in the period $t + 1$.

The timing of events in this predictor is as follows: first, at the beginning of the period agents in both groups have to decide based on their expected profits whether they will build up their capacity to process information or they will decide to free ride and decrease their capacity to process information. This stage is modelled by the NLD. Next, based on this decision and their previous decisions about building up the capacity to process information, it is determined whether any agents are relegated from the rational group to the boundedly rational group or upgraded from the boundedly rational group to the rational group. After determining the proportions of agents in each group, in the last stage we observe which agents in the rational group observe all the necessary information to use the perfect foresight. This at the end determines the proportion of agents that use the perfect foresight and the naive predictor. Graphic exposition of the timing of events can be found in Figure A1 in Appendix.

3.1 Building up Capacity to Process Information and Free Riding

When deciding to build up the capacity to process information or to free ride, one has to compare expected profits from both alternatives. In order to determine them, we have to specify the probabilities of being relegated to the boundedly rational group for agents that are in the rational group and for agents in the boundedly rational group the probability that they actually reach the critical mass of the capacity to process information and are upgraded to the rational group. These will then determine the weights in the expected profits. As mentioned above, it can happen that when rational agents invest into building up the capacity, they are

simply not receiving all the information that is necessary to form rational expectations. We assume that they know the critical level of the capacity to process information; however they do not observe their own capacity to process information, but only the group average. Therefore, they cannot free ride and maintain the level of the capacity that is just above the critical level.⁸ In Section 4.3 we explore an alternative formulation for probabilities of being upgraded and relegated where we assume that agents learn and estimate them from past data on entries and exits.

The probabilities for different groups of agents, before the decision regarding the process of building up the capacity, are defined in Table 2:

	Rational group		Boundedly rational group
	Perfect foresight	Naive rule	Naive rule
Building up capacity	λ	$1 - \lambda$	$1 - \lambda\vartheta_{2,t+1}$
Free riding	$\lambda\vartheta_{1,t+1}$	$1 - \lambda\vartheta_{1,t+1}$	1

Table 2: Probability of using a specific rule

Agents forecast λ , since they only observe realizations of it at the end of the period. So, for example, in the previous period an agent was in the rational group, and if she decides to build up the capacity in this period the probability that she will use the perfect foresight in the current period is λ . Alternatively, if they choose to free ride their probability decreases by $\vartheta_{1,t+1} = 1 - \frac{\log(1-\epsilon)}{\log\left(\frac{\gamma_{0,0}}{\bar{\gamma}_{1,t}^a}\right)}$. $\vartheta_{1,t+1}$ is the probability that they are relegated and it is assumed to equal to the inverse of the number of periods that agents have to free ride in order to be relegated (the derivation of the number of periods they have to free ride comes directly from Proposition 3 below). $\bar{\gamma}_{h,t}^a$ is the "adjusted" capacity to process information⁹ in group h for time t and $\gamma_{0,0}$ is the critical level of the capacity to process information in order to be able to use the perfect foresight. ψ is the gain in the capacity to process information if agents decide to build up capacity, while ϵ is the loss of that capacity if agents free ride. Analogously, $\vartheta_{2,t+1} = \frac{\log(1-\psi)}{\log\left(\frac{\gamma_{0,0}-1}{\bar{\gamma}_{2,t}^a-1}\right)}$ is the probability that agents in the boundedly rational group are upgraded, and is derived from Proposition 4 below.

In this stage agents decide whether to build up the capacity or not based on the expected utility. The appropriate measure for the utility in the cobweb model is profits. Thus, for defining

⁸This might be rational for them from the utility maximization point of view, however this way of behaving might lose its' importance if we introduce new predictors to the model.

⁹Superscript a stands for adjusted for entry and exit and they are defined in equations (15a)-(15b).

the expected utility the net-realized profits for both alternatives have to be computed:

$$\pi_{1,t+1} = (E_1 p_{t+1}) Q_t^* - \frac{1}{2b} (Q_t^*)^2 = \frac{1}{2} b p_{t+1}^2, \quad (9a)$$

$$\pi_{2,t+1} = (E_2 p_{t+1}) Q_t^* - \frac{1}{2b} (Q_t^*)^2 = \frac{1}{2} b p_t (2p_{t+1} - p_t), \quad (9b)$$

where $\pi_{1,t+1}$ ($\pi_{2,t+1}$) is net-realized profits for the perfect foresight (naive predictor) in period $t + 1$ and p_t and p_{t+1} can be replaced by its respective solutions of the form (6).

Results in Table 2 are used to determine expected profits for all groups. They are thus weighted by probabilities of staying in the same group and changing the group. Brock et al. (2006) introduce forward looking behavior in this stage and find significant changes in the nonlinear behavior of the model. The expected average utility for each group of agents is computed by comparing the expected profits and the costs of building up the capacity. The average profits are more appropriate then current profits when deciding to build up the capacity or to free ride, as this is a long term decision.

$$u_{1,1,t+1} = u_{1,1,t} + t^{-1} [(\lambda \pi_{1,t+1} + (1 - \lambda) \pi_{2,t+1} + c) - u_{1,1,t}], \quad (10a)$$

$$u_{1,2,t+1} = u_{1,2,t} + t^{-1} [(\lambda \vartheta_{1,t+1} \pi_{1,t+1} + (1 - \lambda \vartheta_{1,t+1}) \pi_{2,t+1}) - u_{1,2,t}], \quad (10b)$$

$$u_{2,1,t+1} = u_{2,1,t} + t^{-1} [(\lambda \vartheta_{2,t+1} \pi_{1,t+1} + (1 - \lambda \vartheta_{2,t+1}) \pi_{2,t+1} + c) - u_{2,1,t}], \quad (10c)$$

$$u_{2,2,t+1} = u_{2,2,t} + t^{-1} [\pi_{2,t} - u_{2,2,t}], \quad (10d)$$

where $u_{h,k,t}$ are the average profits for the subgroup k of the group h at time t ; in subgroup $k = 1$ agents build up their capacity to process information, while in subgroup $k = 2$ they free ride. c represents the cost of building up the capacity to process information. The average utilities in both groups, $\bar{u}_{h,t+1}$, are:

$$\bar{u}_{1,t+1} = \frac{1}{2} \sum_{k=1}^2 u_{1,k,t+1},$$

$$\bar{u}_{2,t+1} = \frac{1}{2} \sum_{k=1}^2 u_{2,k,t+1}.$$

For the analysis in the next section it is useful to define the difference between both alternatives:

$$F_{1,t+1}(n_t) = u_{1,1} - u_{1,2} = \frac{\lambda b}{2t} \sum_{i=1}^t (1 - \vartheta_{1,i+1}) (p_{i+1} - p_i)^2 + c,$$

$$F_{2,t+1}(n_t) = u_{2,1} - u_{2,2} = \frac{\lambda b}{2t} \sum_{i=1}^t \vartheta_{2,i+1} (p_{i+1} - p_i)^2 + c.$$

Note that we are particularly interested in the two specific cases:

$$\begin{aligned} F_1(1) &= \frac{\lambda b}{2} \left(\frac{\log(1-\epsilon)}{\log(\gamma_{0,0})} \right) \left(\text{var}(p) + \frac{1}{(B+b)^2} \right) + c, \\ F_2(0) &= \frac{\lambda b}{2} \left(\frac{\log(1-\psi)}{\log(1-\gamma_{0,0})} \right) \text{var}(\Delta p) + c, \end{aligned}$$

where Δp is $p_{t+1} - p_t$. The selection criterion for this part of the problem is NLD of Brown and von Neumann (1950) that has been recently discussed in Sandholm (2010) and Waters (2009).¹⁰

The two choice version of NLD is explained in the following equation:

$$\eta_{h,t+1}^a = \frac{\eta_{h,t}^a + \beta (\max \{u_{h,1,t+1} - \bar{u}_{h,t+1}, 0\})}{1 + \beta (\max \{u_{h,1,t+1} - \bar{u}_{h,t+1}, 0\}) + \beta (\max \{u_{h,2,t+1} - \bar{u}_{h,t+1}, 0\})}; \quad h = 1, 2. \quad (12)$$

where superscript a stands for "adjusted" for entry and exit of a specific group, and β measures the intensity of preferences. β equal to infinity corresponds to the neoclassical case of perfect rationality. The latter result is valid for the case when costs for building up the capacity are sufficiently small and do not compensate the gain in profits. Under this predictor selection the best performing strategy has a monotonically increasing fraction of followers. The above version of NLD introduces the intensity of choice parameter in a slightly different way than in Waters (2009).¹¹ He analyzes this predictor in a cobweb model and compares the dynamics with the replicator dynamics and multinomial rule. He also confirms the presence of chaotic behavior, but contrary to Brock and Hommes (1997) the steady state under this setting is a repeller and the model exhibits cycles around the steady state. The only difference between the predictor selection explained in Waters (2009) and the one put forward in this paper is that we propose an alternative introduction of the intensity of choice parameter. Waters (2009) introduces it as an exponent, but in order to keep it simple and a monotonically increasing function of the difference between $\beta(u_{1,t+1} - \bar{u}_{t+1})$ and $\beta(u_{2,t+1} - \bar{u}_{t+1})$ for increasing values of β for any combination of values $(u_{1,t+1} - \bar{u}_{t+1})$, we model it as a product.

The average capacities to process information have the following properties: $\bar{\gamma}_{1,0} > \bar{\gamma}_{2,0}$ and

¹⁰Alternatively we can use the replicator dynamics. The speed of convergence can be controlled by ξ and also with β , which is the intensity of choice parameter. The two choice version of replicator dynamics is of the following form:

$$\eta_{h,t+1}^a = \begin{cases} \eta_{h,t}^a - \left(\frac{2}{\pi}\right) \tan^{-1}(\beta[u_{h,1,t+1} - \bar{u}_{h,t+1}]) (1 - \eta_{h,t}^a) & \text{if } u_{h,1,t+1} \geq \bar{u}_{h,t+1} \\ \left[1 + \left(\frac{2}{\pi}\right) \tan^{-1}(\beta[u_{h,1,t+1} - \bar{u}_{h,t+1}])\right] \eta_{h,t}^a & \text{if } u_{h,1,t+1} < \bar{u}_{h,t+1} \end{cases}; \quad h = 1, 2,$$

where function r takes the following form $r(y) = \left(\frac{2}{\pi}\right) \tan^{-1}(\beta y)$.

¹¹Introducing intensity of choice parameter as in Waters (2009) might have a problematic feature when $u_{1,t+1} - \bar{u}_{t+1} < 1$ as he introduces it as an exponent to the previous expression, i.e. $(u_{1,t+1} - \bar{u}_{t+1})^\beta$. This problem does not emerge in a case with two predictors but in a case with more alternative predictors.

$0 \leq \gamma_h \leq 1$ where $h = 1, 2$. They are defined as:

$$\bar{\gamma}_{1,t+1} = \eta_{1,t+1} [1 - (1 - \bar{\gamma}_{1,t}^a)(1 - \psi)] + (1 - \eta_{1,t+1}) [\bar{\gamma}_{1,t}^a(1 - \epsilon)], \quad (13a)$$

$$\bar{\gamma}_{2,t+1} = \eta_{2,t+1} [1 - (1 - \bar{\gamma}_{2,t}^a)(1 - \psi)] + (1 - \eta_{2,t+1}) [\bar{\gamma}_{2,t}^a(1 - \epsilon)]. \quad (13b)$$

ψ is a gain the agent gets if she chooses to build up capacity, while ϵ is the loss if the agent does not choose to build up the capacity. $\bar{\gamma}_{1,t}^a$ is an adjusted value of the capacity to process information after entry and exit take place, and is explicitly defined below. To ensure the boundedness of $\bar{\gamma}_h$, ϵ and ψ are modeled as relative gains and losses. Therefore, it can be rewritten as $[\bar{\gamma}_{h,t}^a + \psi(1 - \bar{\gamma}_{h,t}^a)]$, so the actual gain is $\psi(1 - \bar{\gamma}_{h,t}^a)$. In this set-up the constant small cost buys less and less capacity to process information when the capacity to process information is increasing. In other words, for a constant gain in the capacity to process information absolute cost increases as capacity increases.

Assumption 2 *The following parameters are in the next ranges: $0 < \gamma_{0,0} < 1$; $0 < \epsilon < (1 - \gamma_{0,0})$ and $0 < \psi < \gamma_{0,0}$.*

The above assumption guarantees that agents cannot shift groups every period and that the threshold level of the capacity to process information is in the appropriate range.

3.2 Entry and Exit Processes

In order to define the actual proportion of agents in a specific group at the end of the period, we have to adjust them for the exit and entry. Agents who have been upgraded from the boundedly rational group enter the subgroup of agents in the rational group that build up capacity. Thus, also the proportions of agents that build up capacity at the end of the period have to be modified by the entry and exit as well as the average capacity to process information in both groups. Agents who left the rational group are degraded to a subgroup of boundedly rational agents who are not building up capacity. The dynamics of entry and exit depend on these proportions. Details about the dynamics of entry and exit processes can be found in Appendix B. Here we just state two propositions about the entry and exit processes.

Proposition 3 *If an agent in the rational group free rides for $\left\lceil \frac{\log \gamma_{0,0}}{\log(1-\epsilon)} \right\rceil$ consecutive periods it will be relegated to the boundedly rational group with probability 1.¹²*

¹²Notation $\lceil x \rceil$ stands for rounding the $x \in \mathbb{R}$ to the nearest highest natural number, thus $\lceil x \rceil \in \mathbb{N}$.

Proof. It is straightforward to solve the following difference equation that describes the evolution of the capacity to process information for the rational group when agents are free riding: $\bar{\gamma}_{1,t} = (1 - \epsilon) \bar{\gamma}_{1,t-1}$. The solution is $\bar{\gamma}_{1,t} = (1 - \epsilon)^t$. Thus, after setting $\bar{\gamma}_{1,t} = \gamma_{0,0}$, we get $t = \frac{\log \gamma_{0,0}}{\log(1-\epsilon)}$ if all agents that start to free ride have full capacity to process information. ■

Proposition 4 If an agent in the boundedly rational group builds up the capacity for $\left\lceil \frac{\log(1-\gamma_{0,0})}{\log(1-\psi)} \right\rceil$ consecutive periods it will be promoted to the rational group with probability 1.

Proof. It is straightforward to solve the following difference equation that describes the evolution of the capacity to process information for the boundedly rational group when agents build up the capacity: $\bar{\gamma}_{2,t} = (1 - \psi) \bar{\gamma}_{2,t-1} + \psi$. The solution is $\bar{\gamma}_{2,t} = 1 - (1 - \psi)^t$. Thus, after setting $\bar{\gamma}_{2,t} = \gamma_{0,0}$ and rearrange, we get $(1 - \psi)^t = 1 - \gamma_{0,0}$. Thus, $t = \frac{\log(1-\gamma_{0,0})}{\log(1-\psi)}$ if they start from $\bar{\gamma}_{2,0} = 0$. ■

Before proceeding we have to define the adjusted proportions of building up capacity after the entry and exit takes place:

$$\eta_{1,t+1}^a = \frac{\Gamma_{1,t+1}}{\Gamma_{1,t+1} + \Gamma_{2,t+1}} = \frac{\eta_{1,t+1} \frac{n_t}{\lambda} + \Xi_{3,t+1}}{\frac{n_t}{\lambda} + \Xi_{3,t+1} - \Gamma_{3,t+1}}, \quad (14a)$$

$$\eta_{2,t+1}^a = \frac{\Xi_{2,t+1}}{\Xi_{1,t+1} + \Xi_{2,t+1}} = \frac{\eta_{2,t+1} \left(1 - \frac{n_t}{\lambda}\right) - \Xi_{3,t+1}}{1 - \frac{n_t}{\lambda} - \Xi_{3,t+1} + \Gamma_{3,t+1}}. \quad (14b)$$

We have to redefine as well $\bar{\gamma}_{h,t+1}$, so that the average capacity to process information is updated for possible entries and exits from the relevant groups.

$$\bar{\gamma}_{1,t+1}^a = \frac{\frac{n_t}{\lambda} \bar{\gamma}_{1,t+1} + \Xi_{3,t+1} \left[\gamma_{0,0} + \frac{\psi}{2} (1 - \gamma_{0,0}) \right] - \Gamma_{3,t+1} \left[\gamma_{0,0} \left(1 - \frac{\epsilon}{2}\right) \right]}{\Gamma_{1,t+1} + \Gamma_{2,t+1}}, \quad (15a)$$

$$\bar{\gamma}_{2,t+1}^a = \frac{\frac{n_t}{\lambda} \bar{\gamma}_{2,t+1} + \Gamma_{3,t+1} \left[\gamma_{0,0} \left(1 - \frac{\epsilon}{2}\right) \right] - \Xi_{3,t+1} \left[\gamma_{0,0} + \frac{\psi}{2} (1 - \gamma_{0,0}) \right]}{\Xi_{1,t+1} + \Xi_{2,t+1}}. \quad (15b)$$

It is assumed that the agents who enter the rational group have on average the capacity to process information equal to $\left[\gamma_{0,0} + \frac{\psi}{2} (1 - \gamma_{0,0}) \right]$. The agents who left this group have on average the capacity to process information equal to $\gamma_{0,0} \left(1 - \frac{\epsilon}{2}\right)$.

4 Analyzing the Predictor Choice Dynamics: Steady States

In order to define the Misspecification Heterogeneous Expectations Equilibrium we have to embed the endogenous determination of n_t in the Heterogeneous Expectations Equilibrium concept defined in Section 2. This procedure is similar to the one in Branch and Evans (2006).

The determination of n_t is detailed in Section 3. Note that the solution of the equation (7), which governs the evolution of n_{t+1} is:

$$n_{t+1} = n_0 + \sum_{s=0}^{t+1} \lambda^{t-s} (\Xi_{3,s} - \Gamma_{3,s}).$$

Therefore, it is straightforward to see that:

Proposition 5 n_{t+1} is in its steady state ($n_1 \rightarrow n^*$) only when $\{\Xi_{3,1}, \Gamma_{3,1}\} \rightarrow \{\Xi_3^*, \Gamma_3^*\}$ and $\Xi_3^* = \Gamma_3^*$.

Ξ_3^* , Γ_3^* and n^* denote the steady states of Ξ_3 , Γ_3 and n , respectively. Therefore the next question is when $\Xi_{3,t}$ and $\Gamma_{3,t}$ converge to their steady states. It is again straightforward to show that only when $\eta_{1,t+1}$ and $\eta_{2,t+1}$ are in their steady states $\{\eta_1^*, \eta_2^*\}$, all the remaining endogenous variables in this system $(\bar{\gamma}_{1,t+1}^a, \bar{\gamma}_{2,t+1}^a, \Xi_{3,t+1}, \Gamma_{3,t+1})$ can settle to their respective long run steady states. In order to find all possible steady states of n_{t+1} we have to determine all possible combinations of steady states for $\eta_{1,t+1}$ and $\eta_{2,t+1}$, $\{\eta_1^*, \eta_2^*\}$.

Corollary 6 n_{t+1} can be in its steady state ($n_1 \rightarrow n^*$) in the long run only when $\{\eta_{1,1}, \eta_{2,1}\} \rightarrow \{\eta_1^*, \eta_2^*\}$.¹³

In general, we cannot rule out multiple steady states in the model. To prove uniqueness, similarly to Branch and Evans (2006), we would have to prove that both $F_{1,t+1}$ and $F_{2,t+1}$ are decreasing in the share of agents that use a perfect foresight for $\alpha > 0$, and that $\bar{\gamma}_{1,t+1}^a$ and $\bar{\gamma}_{2,t+1}^a$ are monotonically increasing in n_t .¹⁴ However, due to our setup this is not feasible. The reason is that the relationship between $\eta_{1,t+1}$ ($\eta_{2,t+1}$) and $F_{1,t+1}$ ($F_{2,t+1}$) is difficult to analyze as there are two effects at work. One is a direct effect from increasing the $F_{1,t+1}$ and $F_{2,t+1}$ via $\eta_{1,t+1}$ and $\eta_{2,t+1}$ on n_{t+1} . Moreover, there is also the effect of the change in the probability of being relegated or upgraded. This effect comes from the change in $\bar{\gamma}_{1,t+1}^a$ ($\bar{\gamma}_{2,t+1}^a$). In the case of the rational group this effect actually goes in the same direction as the direct effect on $F_{1,t+1}$. However, in the case of $F_{2,t+1}$ these effects work in the opposite direction. When more and more agents build up the capacity to process information, the higher is the probability of being

¹³"Long run" is pointed out because in transition periods $\eta_{1,t+1}$ and $\eta_{2,t+1}$ could move away from their steady states while n_t will not react until exit and entry takes place. Note that it might take several periods of, e.g., free riding before the person is actually relegated to the boundedly rational group.

¹⁴From the arguments in Waters (2009) we can deduct that $\eta_{1,t+1}$ and $\eta_{2,t+1}$ are monotonically increasing functions. However, the functions are not globally continuously differentiable due to the construction of the NLD. We can only say that the share of agents that build up capacity is continuous in $F_{1,t+1}$ and $F_{2,t+1}$.

upgraded. Therefore, if this effect dominates the direct effect, we would have the possibility of multiple equilibria in the determination of $\eta_{2,t+1}$.¹⁵

In order for the system to remain in the long-run equilibrium the ratio of agents that build up capacity has to be in a steady state in both groups.¹⁶ The steady states for the rational group are $\eta_1^* = \{0, \hat{\eta}_1^*, 1\}$, where $\hat{\eta}_1^* \in (0, 1)$, and similarly for the boundedly rational group we have $\eta_2^* = \{0, \hat{\eta}_2^*, 1\}$, where $\hat{\eta}_2^* \in (0, 1)$, as can be deduced from the equation (7) and taking into account the entry and exit process. Thus, all combinations of steady states of η_1^* and η_2^* can potentially result in a steady state for the n_t . We analyze these steady states for different values of α , β , and c .

Stability of the steady state depends on the properties of the cobweb model. All steady states are stable for $0 > \alpha > 1$. However, if $\alpha > 1$ then the stability depends on the steady state level of n^* . For $n^* \geq \frac{1}{2}$ the steady state is always stable, otherwise the next condition has to be fulfilled for the steady state to be stable: $\alpha < \frac{1}{1-2n^*}$. For example, the steady state where all agents use the simpler predictor is unstable when $\alpha < 1$. These conditions are deduced from the law of motion for p_{t+1} in equation (5).

4.1 Basic Case

As we will show, in this framework it is possible to obtain two steady states where the dominated predictor vanishes and two types of steady states where there is persistent heterogeneity. An important role in this framework is played by the self referential property of the model itself which, in combination with other parameters, defines the resulting steady state. We first analyze the steady states for the case with finite β and later with infinite values of the preference of choice parameter.

4.1.1 Finite Intensity of Choice Parameter

When β is finite and agents are boundedly rational regarding the decision about building up capacity to process information, we can observe one of the following steady states:

Proposition 7 *Misspecification Heterogeneous Expectations Equilibrium when β is finite and $-c > 0$ has one of the following properties:*

¹⁵This problem could be avoided by assuming constant probabilities of being upgraded and relegated in the fitness functions. Then it is possible to prove that $F_{1,t+1}$ and $F_{2,t+1}$ are strictly decreasing in n_t .

¹⁶In the short run this is only a sufficient but not a necessary condition, as for "short" periods of time $\eta_{h,t+1}$ can vary and n_{t+1} would still be in the equilibrium (entry and exit process).

- $n_1 \rightarrow n^* = \hat{n}^* \in (0, \lambda)$ where $F_1(1) \geq 0$ and $F_2(0) \leq 0$ the equilibrium exhibits heterogeneity.
- $n_1 \rightarrow n^* = \hat{n}^* \in (0, \lambda)$ where $F_1(\hat{\eta}_1^*) = 0$, $F_2(\hat{\eta}_2^*) = 0$, and $\Xi_3^* = \Gamma_3^* \neq 0$ model also exhibits heterogeneity.
- $n_1 \rightarrow n^* = 0$ where $F_1(1) < 0$ and $F_2(0) \leq 0$ the equilibrium exhibits vanishing of agents in the rational group.
- $n_1 \rightarrow n^* = \lambda$ where $F_1(1) \geq 0$ and $F_2(0) > 0$ the equilibrium exhibits vanishing of agents in the boundedly rational group.

In principle, we have nine possible combinations of steady states of $\{\eta_1^*, \eta_2^*\}$, as we mentioned above. However, as explained in the proposition, there are only four plausible different steady states. The steady state mentioned in the first bullet point of Proposition 7 exhibits the intrinsic heterogeneity and occurs when ratios of building up capacity are in the following steady states for the corresponding groups: $\eta_1^* = 1$ and $\eta_2^* = 0$. The sufficient condition for this equilibrium to occur is $\frac{\lambda b}{2} \left(\frac{\log(1-\epsilon)}{\log(\gamma_{0,0})} \right) \left(\text{var}(p) + \frac{1}{(B+b)^2} \right) \geq -c \geq \frac{\lambda b}{2} \left(\frac{\log(1-\psi)}{\log(1-\gamma_{0,0})} \right) \text{var}(\Delta p)$. In this steady state we end up also if $\eta_1^* = 1$ and $\eta_2^* = \hat{\eta}_2^*$. The intuition for this is that some agents are upgraded after some periods if $\eta_2^* = \hat{\eta}_2^*$. Thus, the probability of being upgraded decreases and is not high enough to offset the cost of building up the capacity to process information, and consequently $\eta_{2,1} \rightarrow \eta_2^* = 0$. When β is finite we have an additional interior steady state that manifests the intrinsic heterogeneity (the case when $\Xi_3^* = \Gamma_3^* \neq 0$), i.e., the steady state when the entry and exit processes are equal and positive. The possibility for this steady state is evident in the equation (7), as when $\Xi_3^* = \Gamma_3^*$ the system is in a steady state. This steady state occurs when $\eta_1^* = \hat{\eta}_1^*$ and $\eta_2^* = \hat{\eta}_2^*$ and also when $\eta_1^* = \hat{\eta}_1^*$ and $\eta_2^* = 1$, however this is only a necessary condition. A sufficient condition is that the entry and exit from and to the rational group are equal. Next, we have to prove that the following two steady states could never occur. The case when $\eta_1^* = 0$ and $\eta_2^* = 1$ and the case when $\eta_1^* = 0$ and $\eta_2^* = \hat{\eta}_2^*$ are not plausible, as the condition for $\eta_1^* = 0$ is $\frac{\lambda b}{2} \left(\text{var}(p) + \frac{1}{(B+b)^2} \right) \leq -c$. Note that in this case the probability to be relegated when free riding is 1. Thus, agents are relegated, which then decreases the ratio of agents that build up capacity in the boundedly rational group and that further decreases the $\eta_{2,t}$ as the F_2 decreases. Actually under $\frac{\lambda b}{2} \left(\text{var}(p) + \frac{1}{(B+b)^2} \right) \leq -c$ agents in the boundedly rational group would never even start to build up their capacity as

$\frac{\lambda b}{2} \left(\text{var}(p) + \frac{1}{(B+b)^2} \right) > \frac{\lambda b}{2} \left(\frac{\log(1-\psi)}{\log(1-\gamma_{0,0})} \right) \text{var}(\Delta p)$. Therefore, these two cases would end up in the steady state where all agents are in the boundedly rational group. This state also occurs when $\eta_1^* = 0$ and $\eta_2^* = 0$, and when $\eta_1^* = \hat{\eta}_1^*$ and $\eta_2^* = 0$. For the former the condition is $\frac{\lambda b}{2} \left(\frac{\log(1-\psi)}{\log(1-\gamma_{0,0})} \right) \text{var}(\Delta p) < -c$. The latter also converges to the same steady state although $\eta_1^* = \hat{\eta}_1^*$ as the constant proportion of rational agents is relegated until all agents resort to the naive rule. The steady state when all agents are in the rational group arise in the case when $\eta_1^* = 1$ and $\eta_2^* = 1$. The condition for this steady state is $\frac{\lambda b}{2} \left(\frac{\log(1-\epsilon)}{\log(\gamma_{0,0})} \right) \left(\text{var}(p) + \frac{1}{(B+b)^2} \right) > -c$. All steady states mentioned in the proposition above can arise for all values of α that satisfy Assumption 1.

Proposition 8 *Misspecification Heterogeneous Expectations Equilibrium when β is finite and $c = 0$ has the following property:*

- $n_1 \rightarrow n^* = \lambda$ where $F_1(1) > 0$ and $F_2(0) > 0$ the equilibrium exhibits vanishing of agents in the boundedly rational group.

When there is no cost for building up the capacity to process information, then all agents move to the rational group. This result is different from the one in Branch and Evans (2006), as in their case the predictors exhibit identical performance in the long run. Therefore, when they are both costless the heterogeneity can arise as well. However, if we would have in our framework two rules that have an identical performance in the long run and are equally costly, then we would also observe the intrinsic heterogeneity between these two rules.

Vanishing of the strictly dominating predictor has been previously observed under finite β only in the case of replicator dynamics. Also in our case when all agents mass on the less sophisticated predictor this steady state is unstable as previously pointed out in Brock and Hommes (1997) and Branch and McGough (2008) for $\alpha > 1$. If the latter condition is not satisfied, this steady state is stable. It is important that the predictor allows also for the increasing fractions of the strictly dominating predictor. This is one of the two essential properties discussed in Waters (2009), and it is also supported by the empirical literature. In this case we could say that β governs the speed of convergence to the steady state.

When agents are boundedly rational regarding their decision to build up the capacity or free ride we have two different intrinsic heterogeneous steady states: where all agents in the rational group build up the capacity and where all agents who are in the boundedly rational group

free ride and when there is positive entry and exit, $\Xi_3^* = \Gamma_3^* \neq 0$. The intrinsic heterogeneity under finite β is only observed in the Brock and Hommes (1997) setup with a multinomial predictor which is inherently stochastic. Otherwise, intrinsic heterogeneity is a product of the self referential property of the cobweb model and the structure of the predictor choice dynamics. To compare the resulting intrinsic heterogeneity of the system with the one of Branch and Evans (2006), we can observe that the results are quite different. They show that heterogeneity can occur under the choice of forecasting models that produce a similar performance in the long run. To the contrary, following results from the experimental and survey data literature, this paper shows that the heterogeneity is not just a feature of agents with equal abilities, but also of agents with different abilities. This is a new perspective of possible roots of heterogeneity observed in the data.

We also briefly outline the results in a case when the $\eta_{1,t+1}$ and $\eta_{2,t+1}$ evolve according to the Brock and Hommes (1997) predictor. When β is finite and $c < 0$ there is only one possible steady state and that is when $\eta_1^* = \hat{\eta}_1^*$ and $\eta_2^* = \hat{\eta}_2^*$. Therefore, the only steady state of n_t occurs when $\Xi_3^* = \Gamma_3^*$. This is the main reason why we choose either NLD or replicator dynamics for our baseline case as we do not want to restrict the predictor choice dynamics to have only the steady state where there is heterogeneity. We assert that it is important that also when agents are not perfectly rational in deciding about building up the capacity to process information, the model exhibits the possibility of converging to homogeneous expectations.

Finally, also the role of λ needs to be discussed. If λ is below the minimum level of agents using perfect foresight that guarantees stability of the steady state then it can have implications for the dynamics of the model. Otherwise, it does not play a role for the main results in the paper and it just caps $n^* \leq \lambda$ in the the steady state where all agents are in the rational group as long as the Assumption 2 is fulfilled. If we would set $\lambda = 1$ then all agents in the rational group would be using perfect foresight and all agents in the boundedly rational group the naive predictor. ψ and ϵ also have an impact on the regions of parameters where we observe particular steady states. They influence profits for different alternatives and thus through F_1 and F_2 affect the regions of c where we observe a particular steady state of n_t . For example, if we increase ψ this would, ceteris paribus, decrease the region of c where we observe intrinsic heterogeneity (where $\Xi_3^* = \Gamma_3^* = 0$)¹⁷ and decrease the region of c where we observe $n_1 \rightarrow n^* = 0$, where all

¹⁷The effect on the intrinsic heterogeneous steady state where $\Xi_3^* = \Gamma_3^* \neq 0$ depends on the levels of $\bar{\gamma}_{1,t+1}^a$ and $\bar{\gamma}_{2,t+1}^a$.

agents are in the rational group. It would, however, increase the region of c where we observe $n_1 \rightarrow n^* = \lambda$ and all agents are in the boundedly rational group.

4.1.2 Infinite Intensity of Choice Parameter – the Neoclassical Case

Under infinite β , we do not observe the convergence to the intrinsic heterogeneous steady state where there are positive values for entry and exit, but we can still observe the convergence to the other steady state where there is intrinsic heterogeneity. It is interesting that also in this perfectly rational case the generality of the model is still present. Depending on the parameterization, both massing on the strictly dominating predictor and intrinsic heterogeneity are possible.

Definition 9 *A model is said to exhibit intrinsic heterogeneity if Misspecification Heterogeneous Expectations Equilibrium exists for all $\beta > 0$ and there exist $\bar{n} < \lambda$ such that $n^* \leq \bar{n}$, for all β .*

Condition 10 *As $\beta \rightarrow \infty$, $n_1 \rightarrow n^* = \hat{n}^* \in (0, \lambda)$ where $F_1(1) \geq 0$ and $F_2(0) \leq 0$.*

Proposition 11 *Assume that the above condition is fulfilled, then the model exhibits long run intrinsic heterogeneity.*

More generally, we can show that the MHEE when $\beta \rightarrow \infty$ has one of the properties stated in the next proposition.

Proposition 12 *Misspecification Heterogeneous Expectations Equilibrium when $\beta \rightarrow \infty$ and $-c > 0$ has one of the following properties:*

- $n_1 \rightarrow n^* = \hat{n}^* \in (0, \lambda)$ where $F_1(1) \geq 0$ and $F_2(0) \leq 0$ the equilibrium exhibits intrinsic heterogeneity.
- $n_1 \rightarrow n^* = 0$ where $F_1(1) < 0$ and $F_2(0) \leq 0$ the equilibrium exhibits vanishing of agents in the rational group.
- $n_1 \rightarrow n^* = \lambda$ where $F_1(1) \geq 0$ and $F_2(0) > 0$ the equilibrium exhibits vanishing of agents in the boundedly rational group.

Remark 13 *The fourth (theoretically possible) case where $F_1(0) \leq 0$ and $F_2(1) \geq 0$ never occurs as agents who build up capacity in the boundedly rational group are promoted to the rational group.*

This condition asserts that there is no movement between the two groups. Therefore, the intrinsic heterogeneous steady state can emerge only for the steady state where there is no movement between groups. In order that the above condition is fulfilled, α , the parameter that describes the self referential property of the model, has to be sufficiently high and c has to be within the appropriate range. Generally, when $\beta \rightarrow \infty$ we have four different possible combinations of steady states: $\{\eta_1^*, \eta_2^*\}$. Compared to the previous subsection when β is finite, all agents in one group mass on the dominating predictor. Therefore the steady state for the ratio of agents that build up capacity is either 0 or 1. As before, the case $\eta_1^* = 0$ and $\eta_2^* = 1$ is not plausible. The first bullet point of Proposition 12 corresponds to $\eta_1^* = 1$ and $\eta_2^* = 0$, while the second bullet point coincides with $\eta_1^* = 0$ and $\eta_2^* = 0$. The third bullet point is relevant when all agents in the economy are building up capacity, i.e., $\eta_1^* = 1$ and $\eta_2^* = 1$.

Proposition 14 *Misspecification Heterogeneous Expectations Equilibrium when $\beta \rightarrow \infty$ and $c = 0$ has the following property:*

- $n_1 \rightarrow n^* = \lambda$ where $F_1(1) \geq 0$ and $F_2(0) > 0$ the equilibrium exhibits vanishing of agents in the boundedly rational group.

When there is no cost for building up the capacity to process information, all agents enter the rational group and try to use perfect foresight, as it is strictly dominating in our framework. In order that heterogeneity arises in our framework with infinite β , a rule has to be strictly dominating only for one group, and not for both.

Under infinite β , where agents rationally decide whether to invest into building up the capacity to process information or free riding, we observe that in certain circumstances the strictly dominated predictor vanishes. A similar result is obtained in Brock and Hommes (1997) and Branch and McGough (2008), and it is a quite standard result that heterogeneity is not present in this neoclassical case.

An interesting result of our paper is the occurrence of intrinsic heterogeneity, also in the neoclassical setup where agents rationally decide about building up capacity. This result was previously obtained in Branch and Evans (2006), however the choice of rules in their framework is quite different, as agents decide between two underparametrized rules which exhibit identical performance in the long run. We obtain this result under a different setting where agents choose among perfect foresight and myopic predictors. These predictors do not exhibit an identical performance, and furthermore, using the perfect foresight is not costless. This result is closer

to the behavior observed across different demographic groups in the survey data literature and in laboratory experiments.

Qualitatively, there are no differences between the results obtained above for the neoclassical case when $\beta \rightarrow \infty$, and the results that would be obtained with the Brock and Hommes (1997) predictor for the choice of building up the capacity to process information. In the case of the Brock and Hommes (1997) predictor we would get exactly the same three possible steady states as in the proposition above. This gives another dimension to our results, as in the perfectly rational case we could use any existing predictor known in the literature and our results would still exhibit very similar characteristics.

4.1.3 Some Numerical Exercises

To study the relevance of parameters of interest on the realizations of different steady states we perform several numerical exercises. In particular, we are interested how the self referential property and the intensity of choice parameter influence the regions of c where we observe a specific steady state of n_t . The parameters that are varied in different cases are thus c , B , and β . Basic parameters used in numerical exercises (which do not vary in different cases) are summarized in Table 3.

$\psi = 0.02$	$\lambda = 0.9$	$\gamma_{0,0} = 0.75$
$\epsilon = 0.01$	$b = 1.35$	

Table 3: Parameter values

In the following figures we focus on the bifurcation diagrams for c . For each value of c 500 realization (periods 12,501 – 13,000) of n_t (after 12,500 transition periods)¹⁸ are plotted. Thus, if we observe a single realization of n_t (small diamond in a Figure at a specific value of c) for all 500 periods it means that the system has settled to a steady state for a particular value of c . If we would observe a region of values for n_t at a specific c then this would suggest that n_t has not settled (yet) to a steady state. In Figure 1 we can observe the importance of the self referential property of the cobweb model. These diagrams are calculated by assuming $\beta = 1$ and in the space for $c \in [-0.02, 0]$. If $B = 0.5$ ($\alpha = 2.7$), as in Figure 1(a), we can see that for c , smaller than approximately -0.006 , the model results in an intrinsic heterogeneous steady state. To be specific, in this case we can observe the intrinsic heterogeneous steady state where all agents in

¹⁸A long transition period is needed due to a convergence process that is relatively slow, as our model is defined in two stages (and in the first stage the average profits have to converge).

the rational group build up capacity and all agents in the boundedly rational group free ride. The ratio of agents using the perfect foresight (n_t) increases when c increases. For c between -0.006 and 0 the steady state where all agents are in the rational group prevails.

Figure 1(a)-(d): Bifurcation diagrams for c with $\beta = 1$.

In Figure 1(b) we repeat the same exercise with $B = 0.9$ ($\alpha = 1.5$). Immediately, we can see that the range of values of c for which we observe the steady state where all agents are in the rational group decreases significantly on the right hand side. Only for c higher than approximately -0.0042 , all agents are in the rational group. However, all agents are in the rational group also for c lower than -0.015 . Note that this steady state is not stable and does not represent a "long-run" steady state, it is just temporary as agents will start to free ride and thus some will be moved to the boundedly rational group.¹⁹ For all the other values of c in the

¹⁹This process is very slow under current parametrization and 500 periods for which these plots are made might not be enough to see this dynamics.

relevant range, the model ends up in an intrinsic heterogeneous steady state where all agents in the rational group build up their capacity and all agents in the boundedly rational group free ride; thus there is no movement between groups. Compared to previous diagrams, in Figure 1(c) we further decrease the self referential property of the model. The area where we observe intrinsic heterogeneous steady states further shrinks. The model converges to homogeneous expectations where all agents are in the rational group if $c < -0.011$, while the area with the same characteristics on the right hand side of the figure becomes again smaller. In the last exercise, Figure 1(d), we choose $B = 1.7$, so that α is less than 1 ($\alpha \approx 0.79$). In this case the properties of the model are quite different, because the steady state where all agents use the less sophisticated rule is stable. Thus, for the majority of the relevant space of c , we observe the convergence towards homogeneous expectations, where they all use the naive predictor. The area with intrinsic heterogeneous expectations again becomes smaller. Because α is less than 1, we do not observe the steady state where all agents are in the rational group on the left hand side of this figure. In Figures 1 it is also possible to observe for which values the system is stable. Comparing Figures 1(b) and (c) it is possible to see that in Figure 1(c) the system is stable for lower values of n^* .

Figure 2 repeats the exercises reported above with $\beta = 10^{50}$. Thus, we try to examine the influence of the intensity of choice parameter for the determination of the steady state. Qualitatively, the results are quite similar to those obtained in the previous exercise. With a high value of β we observe only three values for $n^* = \{0, 0.6, 0.9\}$ in this exercise. In Figure 2(a) we observe only regions where intrinsic heterogeneous steady state prevails and the steady state where all agents are in the rational group ($c < -0.08$). In other figures there exist another region where the system might, for a short time, converge to the steady state where all agents are in the rational group. Nevertheless, after some time the average volatility of price decreases, and thus some agents decide to free ride and are relegated to the boundedly rational group. As the new equilibrium reaches a level of the ratio of agents using perfect foresight (n_t) for which the steady state is unstable, the price volatility increases and the system is pushed again to the equilibrium where all agents are in the rational group. This process repeats continuously, and for higher c we generally observe more frequent switches between the steady states.

Figure 2(a)-(d): Bifurcation diagrams for c with $\beta = 10^{50}$.

4.2 Constant Gain Updating of Fitness Functions

In this subsection we test the relevance of the assumption that investing into capacity to process information is a long term decision. We replace equations that explain the determination of average profits (10a)-(10d) with the following equations:

$$u_{1,1,t+1} = u_{1,1,t} + \varrho [(\lambda\pi_{1,t+1} + (1 - \lambda)\pi_{2,t+1} + c) - u_{1,1,t}], \quad (16a)$$

$$u_{1,2,t+1} = u_{1,2,t} + \varrho [(\lambda\vartheta_{1,t+1}\pi_{1,t+1} + (1 - \lambda\vartheta_{1,t+1})\pi_{2,t+1}) - u_{1,2,t}], \quad (16b)$$

$$u_{2,1,t+1} = u_{2,1,t} + \varrho [(\lambda\vartheta_{2,t+1}\pi_{1,t+1} + (1 - \lambda\vartheta_{2,t+1})\pi_{2,t+1} + c) - u_{2,1,t}], \quad (16c)$$

$$u_{2,2,t+1} = u_{2,2,t} + \varrho [\pi_{2,t} - u_{2,2,t}], \quad (16d)$$

where ϱ is a constant gain. When $\varrho = 1$ only the most recent expected profits matter, as in Brock and Hommes (1997), and when $\varrho \rightarrow 0$ this estimator converges to the limiting case of average profits which is analyzed in the previous subsection. The results regarding steady states

in this section are the same as those obtained in the basic case, as long as ϱ is sufficiently small. When ϱ is relatively high and agents only take into consideration last few profits, variability of n_t around its steady state emerges. In some cases, if ϱ is close to 1 and if the relevant steady state exhibits the intrinsic heterogeneity, the evolution of n_t might resemble switching between different steady states in a cyclical manner. Under this estimator of alternative profits, if β increases, the volatility further increases. A numerical exercise with a geometric estimator is presented in Figure A2 in the Appendix.

In this case we can show the generality of the model. If we set in $\varrho = 1$ and we disregard the Assumption 2 and instead assuming that $\epsilon > \gamma_{0,0}$ and that $\psi > 1 - \gamma_{0,0}$ then our model collapses to the predictor that is used in the second stage. Therefore, we allow agents to switch between predictors every period based on the most recent expected profits. In case we use NLD in the second stage it collapses to the stochastic version of the model analyzed in Waters (2009).

4.3 Learning about the Probabilities of Switching to the Other Group

As we learned above, the probabilities of being upgraded or relegated play a crucial role in the predictor choice dynamics. In the basic case we assume that agents only observe the average capacity to process information in the group and the threshold level. Based on this information they calculate the average probability of being relegated or promoted to a specific group. In this section we assume that agents learn from the actual past data on the proportion of agents that were relegated. We assume sample average learning. Thus, we replace $\vartheta_{1,t+1}$ and $\vartheta_{2,t+1}$ in (10b)-(10c) with:

$$\begin{aligned}\vartheta_{1,t+1} &= \vartheta_{1,t} + t^{-1} [(x_{1,t} + \chi) - \vartheta_{1,t}], \\ \vartheta_{2,t+1} &= \vartheta_{2,t} + t^{-1} [(x_{2,t} + \chi) - \vartheta_{2,t}],\end{aligned}$$

where χ is a small positive number ($0 < \chi < 1$). We have to introduce this constant as otherwise it might occur that agents in the boundedly rational group would never have the opportunity to build up their capacity when they are all free riding. Therefore, we have to prevent that the probability becomes asymptotically close to 0. Also, boundedly rational agents have to have the opportunity to start learning and building up their capacity.

If we compare the results with our basic case, we can immediately state that for the most combinations of parameter values we get increased volatility of n_t when agents learn about the probability of being upgraded (relegated). In this case we could say that agents are engaged

in the double learning as they learn about the probabilities and about the alternative expected profits. Results about the steady states are quite similar to those obtained in our basic case. However, we have to point out that for $c \neq 0$, the steady state where all agents are in the rational group could not always be stable in the long run.²⁰ This is simulated in Figure A3 in the Appendix. The steady state that exhibits the intrinsic heterogeneity can be arbitrarily close to the steady state where everyone is in the rational group. Even more, the model can be at the latter steady state for a certain period of time before converging to the intrinsic heterogeneous steady state. Nevertheless, when the probability of being relegated becomes too close to 0, so it could not offset the cost of using the perfect foresight, the model would depart from that steady state. This result is robust to the value of the intensity of choice parameter. However, all steady states outlined above can occur in this setup and can be stable in the long run. If sample average learning would be replaced with the geometric estimator of the probabilities, the variability of n_t would further increase.

5 Conclusion

This paper presents a new predictor choice dynamics that combines the properties of Adaptively Rational Expectations Dynamics and Replicator Equilibrium Dynamics. It generalizes the main result from Branch and Evans (2006) on the intrinsic heterogeneity. We show that it is possible to obtain the intrinsic heterogeneity also between inherently different predictors and not only between predictors that exhibit identical performance in the long run. This result is obtained by modeling the predictor choice dynamics based on several potential roots of heterogeneity that are usually observed in the data. We particularly focus on the fact that not all agents can use sophisticated predictors (perfect foresight) if they do not have high enough capacity to process information. They have to build up capacity for several periods before they can use the rational predictor. Furthermore, we model this decision as a long term one where agents decide based on the average expected difference in profits. We also argue that information acquisition is a short term decision and that we have to separate these two processes to efficiently analyze the importance of different roots of heterogeneity.

The resulting predictor is determined in two stages. Agents in the first stage choose between the costly building up of the capacity and free riding. This stage is modeled by adopting the

²⁰These results depend on a particular parametrization of the model and χ .

New Learning Dynamics as in Waters (2009). After this stage, if agents have the capacity to process information greater than the critical level, then they are in the rational group. Whether they can use a sophisticated predictor depends also on the information acquisition. Only in the case when they observe all the relevant information, they can use the rational predictor. So even when agents build up the capacity it can happen that they might not be able to forecast with the perfect foresight.

The resulting predictor has several advantages as it allows for homogeneous and heterogeneous expectations in a steady state regardless of the value of the intensity of choice parameter. In previous studies this parameter has important implications for the resulting steady state and only Branch and Evans (2006) show that they can obtain coexistence of different predictors in a steady state in a neoclassical example when the intensity of choice parameter is equal to infinity. Moreover, in our approach two different scenarios lead to the intrinsic heterogeneity. The first one occurs when all agents in the rational group build up capacity and when all agents who use the simpler predictor free ride. In this case there is no switching between the groups. However, there is another steady state that is possible in our framework, where the number of agents that are relegated from the rational group is equal to the number of agents that are upgraded from the boundedly rational group. We also confirm one of the main results advanced by Brock and Hommes (1997) and Branch and McGough (2008), i.e., the steady state attained when all agents rely on a naive rule is unstable under a sufficiently large ratio of slopes of the demand and the supply.

References

- Anderson, R. D. J., Becker, R., and Osborn, D. R. (2010). Heterogeneity in consumers’ learning about inflation. Newcastle Discussion Papers in Economics 2010/02, Newcastle University.
- Anufriev, M. and Hommes, C. (2009). Evolutionary selection of individual expectations and aggregate outcomes. CeNDEF Working Papers 09-09, Universiteit van Amsterdam.
- Assenza, T., Heemeijer, P., Hommes, C., and Massaro, D. (2011). Individual expectations and aggregate macro behavior. CeNDEF Working Papers 11-01, Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.
- Ball, L., Mankiw, G. N., and Reis, R. (2005). Monetary policy for inattentive economies. *Journal of Monetary Economics*, 52(4):703–725.
- Branch, W., Carlson, J., Evans, G., and McGough, B. (2009). Monetary policy, endogenous inattention and the volatility trade-off. *Economic Journal*, 119(534):123–157.
- Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal*, 114(497):592–621.
- Branch, W. A. (2007). Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control*, 31(1):245–276.
- Branch, W. A. and Evans, G. W. (2006). Intrinsic heterogeneity in expectation formation. *Journal of Economic Theory*, 127(1):264–295.
- Branch, W. A. and Evans, G. W. (2007). Model uncertainty and endogenous volatility. *Review of Economic Dynamics*, 10(2):207–237.
- Branch, W. A. and McGough, B. (2008). Replicator dynamics in a cobweb model with rationally heterogeneous expectations. *Journal of Economic Behavior and Organization*, 65(2):224–244.
- Branch, W. A. and McGough, B. (2009). A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 33(5):1036–1051.
- Brock, W. A., Dindo, P., and Hommes, C. H. (2006). Adaptive rational equilibrium with forward looking agents. *International Journal of Economic Theory*, 2(3-4):241–278.

- Brock, W. A. and Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65(5):1059–1096.
- Brown, G. W. and von Neumann, J. (1950). Solutions of games by differential equation. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games I, Annals of Mathematics Studies 24*. Princeton University Press, Princeton.
- Bryan, M. F. and Palmqvist, S. (2005). Testing near-rationality using detailed survey data. Working Paper Series 183, Sveriges Riksbank.
- Bryan, M. F. and Venkatu, G. (2001). The demographics of inflation opinion surveys. *Economic Commentary*, (Oct 15).
- Burke, M. A. and Manz, M. (2011). Economic literacy and inflation expectations: evidence from a laboratory experiment. Public Policy Discussion Paper 11-8, Federal Reserve Bank of Boston.
- Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, 118(1):269–298.
- Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116 – 159.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–60.
- Dräger, L. (2011). Inflation perceptions and expectations in sweden - are media reports the ‘missing link’? Macroeconomics and Finance Series 201101, Hamburg University.
- Evans, G. W. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.
- Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). Iq and stock market participation. *Journal of Finance*, 66(6):2121–2164.
- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

- Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control*, 35(1):1–24.
- Hommes, C., Huang, H., and Wang, D. (2005a). A robust rational route to randomness in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 29(6):1043–1072.
- Hommes, C., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2005b). Coordination of expectations in asset pricing experiments. *Review of Financial Studies*, 18(3):955–980.
- Jonung, L. (1981). Perceived and expected rates of inflation in sweden. *American Economic Review*, 71(5):961–68.
- Mankiw, N. G. and Reis, R. (2001). Sticky information: A model of monetary nonneutrality and structural slumps. NBER Working Papers 8614, National Bureau of Economic Research, Inc.
- Mankiw, N. G., Reis, R., and Wolfers, J. (2004). Disagreement about inflation expectations. *NBER Macroeconomics Annual 2003*, 18:209–248.
- Pfajfar, D. and Santoro, E. (2009). Asymmetries in inflation expectations across sociodemographic groups. Mimeo, Tilburg University.
- Pfajfar, D. and Santoro, E. (2010). Heterogeneity, learning and information stickiness in inflation expectations. *Journal of Economic Behavior & Organization*, 75(3):426–444.
- Pfajfar, D. and Santoro, E. (2012). News on inflation and the epidemiology of inflation expectations. *Journal of Money, Credit and Banking*, forthcoming.
- Pfajfar, D. and Žakelj, B. (2012). Experimental evidence on inflation expectation formation. Technical report, Tilburg University, Center for Economic Research.
- Sandholm, W. H. (2010). Pairwise comparison dynamics and evolutionary foundations for nash equilibrium. *Games*, 1(1):3–17.
- Schuster, P. and Sigmund, K. (1983). Replicator dynamics. *Journal of Theoretical Biology*, 100(3):533–538.
- Souleles, N. S. (2004). Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the michigan consumer sentiment surveys. *Journal of Money, Credit and Banking*, 36(1):39–72.

Waters, G. A. (2009). Chaos in the cobweb model with a new learning dynamic. *Journal of Economic Dynamics and Control*, 33(6):1201–1216.

Whalley, J. (2005). Rationality, irrationality and economic cognition. CESifo Working Paper Series CESifo Working Paper No. 1445, CESifo GmbH.

A Additional Figures

Figure A1: Decision process

Figure A2: Simulations with constant gain updating $B = 0.9$, $\beta = 1$, and $c = -0.07$

Figure A3: Simulation with learning about the probabilities of changing the group where

$$B = 0.9, \beta = 1000, \text{ and } c = -0.00003$$

B Entry and Exit processes

Using the properties of $\Gamma_{2,t+1}$ and $\Xi_{2,t+1}$ we can define the number of agents that are relegated to the boundedly rational group ($\Gamma_{3,t+1}$) and the number of agents that are upgraded to the rational group from the boundedly rational group ($\Xi_{3,t+1}$). We first focus on the relegation from the rational group to the boundedly rational group. To draw intuition, let's start with a special case. We know that, in order to be relegated from the rational to the boundedly rational group, agents need to free ride for $s_1 = \frac{\log \gamma_{0,0}}{\log(1-\epsilon)}$ consecutive periods if they start from the full capacity to process information (as stated in the Proposition 3). Thus, if the $\Gamma_{2,t+1}$ is monotonically increasing, agents that have $\frac{\log \gamma_{0,0}}{\log(1-\epsilon)}$ periods ago decided to start free riding, then they will be relegated in the current period. In this case, the number of agents that free ride in the rational group and are consequently relegated in the current period is simply determined as:

$$\Gamma_{3,t+1} = \Gamma_{2,t-s_1+1} - \sum_{i=t-s_2+2}^t \Gamma_{3,i}.$$

In this case $\Gamma_{3,t+1}$ equals to the lowest number of agents that free rode in the time span from $t - s_1$ periods ago up to t adjusted for the exists in the consecutive periods. Here, because of the monotonicity assumption, this equals to the "adjusted" the number of agents that free rode in period $t - s_1 + 1$. However, the procedure for finding the appropriate proportion of agents that exit the group needs to take into account also the possibility that the share of agents does not increase monotonically in this group. Intuitively, if one agent is updating his capacity to

process information for some periods, this has to be compensated by additional periods of free riding before they are relegated. Thus, we have to define this process more generally. We start by fixing t^* whenever there are no agents in group Γ_2 , i.e., when everybody in the rational group builds up capacity and the average capacity to process information in the rational group reaches 1.

$$\Gamma_{3,t+1} = \begin{cases} \min_{c_1} \left\{ \left(\Gamma_{2,t^*} - \sum_{i=t^*+1}^t \Gamma_{3,i} \right), \left(\Gamma_{2,t^*+1} - \sum_{i=t^*+2}^t \Gamma_{3,i} \right), \dots, \Gamma_{2,t} \right\} & \text{if } j_t + 2 = t + 1 \\ 0 & \text{if } otherwise \end{cases},$$

$$j_t = \left\lceil s_1 + \frac{\epsilon \tilde{\gamma}_{1,t}}{\psi(1 - \tilde{\gamma}_{1,t})} r \right\rceil; \quad \tilde{\gamma}_{1,t} = \frac{1}{t - t^*} \sum_{i=t^*}^t \tilde{\gamma}_{1,i},$$

As in the special case above, all ratios have to be adjusted for all consecutive entries and exits when we compare the values of Γ_2 between period t^* and t . We know that $t^* \leq t - s_1 + 1$. In this case we do not search for the highest element in the set, but for the c_1 -th highest element in that set. $c_1 (= t - j_t + 1 + r)$ and r (number of periods that agents build up the capacity) are sequences of natural numbers. We have to ceil $s_1 + \frac{\epsilon \tilde{\gamma}_{1,t}}{\psi(1 - \tilde{\gamma}_{1,t})} r$ to the nearest highest natural number to get j_t .²¹ The notation $\min_{c_1}$ represents the c_1 -th lowest element in that set. Intuitively, in this case we know that some agents have been starting again to build up capacity and thus they need to free ride for j_t periods in total, with r periods of building up capacity in order to be relegated in the next period.

The second issue is to determine the number of agents that are upgraded to the rational group from the boundedly rational group ($\Xi_{3,t+1}$). We determine the proportion of agents that are upgraded to the rational group by analyzing properties of $\Xi_{2,t+1}$. Starting again with a special case: an agent that aims at being promoted to the rational group has to build up the capacity for $\frac{\log(1 - \gamma_{0,0})}{\log(1 - \psi)}$ consecutive periods if her capacity to process information is 0 at the beginning of the process. If we assume that the ratio of agents that are building up the capacity to process information in the boundedly rational group ($\Xi_{2,t+1}$) is monotonically increasing in the relevant interval, then we can define the number of agents that are promotes as:

$$\Xi_{3,t+1} = \Xi_{2,t-s_2+1} - \sum_{i=t-s_2+2}^t \Xi_{3,i}.$$

$\Xi_{3,t+1}$ is in this case equals to the lowest number of agents that build up capacity in the time

²¹In numerical simulations we further simplify the expression $\frac{\epsilon \tilde{\gamma}_{1,t}}{\psi(1 - \tilde{\gamma}_{1,t})}$ by assuming that the ratio of loss to the gain in capacity to process information remains constant at its average, $\frac{\epsilon(1 + \gamma_{0,0})}{4\psi(1 - \gamma_{0,0})}$.

span from $t - s_1$ periods ago up to t adjusted for the exists from the boundedly rational group in the consecutive periods. As we impose the monotonicity of $\Xi_{2,t+1}$, this equals to the "adjusted" the number of agents that build up capacity in period $t - s_1 + 1$. If this assumption is not fulfilled, we have to take into account the possibility that some agents have been free riding for some periods during the process of building up capacity. Again we have to determine the point in time ($t^{**} \leq t - s_2 + 1$) when all agents in the boundedly rational group free ride and their average capacity to process information is $\bar{\gamma}_{2,t} = 0$. In this case, we are not searching for the lowest element in the interval, but for the c_2 -th lowest element in the relevant interval. Thus, the $\Xi_{3,t+1}$ is defined as follows:

$$\Xi_{3,t+1} = \begin{cases} \min_{c_2} \left\{ \left(\Xi_{2,t^{**}} - \sum_{i=t^{**}+1}^t \Xi_{3,i} \right), \left(\Xi_{2,t^{**}+1} - \sum_{i=t^{**}+2}^t \Xi_{3,i} \right), \dots, \Xi_{2,t} \right\} & \text{if } i_t + 2 = t + 1 \\ 0 & \text{if } otherwise \end{cases},$$

$$i_t = \left\lceil s_2 + \frac{\psi(1 - \tilde{\gamma}_{2,t})}{\epsilon \tilde{\gamma}_{2,t}} d \right\rceil; \quad \tilde{\gamma}_{2,t} = \frac{1}{t - t^{**}} \sum_{i=t^{**}}^t \bar{\gamma}_{2,i},$$

where d and c_2 are sequences of natural numbers. The former represents the number of periods that agents free ride and the latter can be defined as follows: $c_2 = t - i_t + 1 + d$. Agents that free ride for d periods have to build up capacity for i_t periods in order to be promoted in the next period.²²

²²In numerical simulations we further simplify the expression $\frac{\psi(1 - \tilde{\gamma}_{2,t})}{\epsilon \tilde{\gamma}_{2,t}}$ by assuming that the ratio of loss compared to the gain in the capacity to process information remains constant at its average, $\frac{4\psi(\frac{1}{2} - \gamma_{0,0})}{\epsilon \gamma_{0,0}}$.