Genetic Algorithm Learning in a New Keynesian Macroeconomic Setup
Working paper — version January 2015

Cars Hommes[a,b], Tomasz Makarewicz[a,b], Domenico Massaro[a,b], and Tom Smits[c]

[a]CeNDEF, University of Amsterdam
[b]Tinbergen Institute, Amsterdam
[c]SEO Economic Research

January 8, 2015

Abstract

In order to understand heterogeneous behaviour amongst agents, empirical data from Learning-to-Forecast (LtF) experiments can be used to construct learning models. This paper follows up on Assenza et al. (2013) by using a genetic algorithms (GA) model to replicate the results from their LtF experiment. In this GA model individuals optimise an adaptive, a trend following and an anchor coefficient in a population of general prediction heuristics. We replicate experimental treatments in a New-Keynesian environment with increasing complexity and use Monte Carlo simulations to investigate how well the model explains the experimental data. We find that the model is able to replicate the three different types of behaviour in the treatments using one GA model. The research furthermore shows that heterogeneous behaviour can be explained by an adaptive, anchor and trend extrapolating component and therewith contributes to the existing literature in the way that GA can be used to explain heterogeneous behaviour in LtF experiments with different types of complexity.

1 Introduction

In this paper we study a simple New Keynesian economy, in which the individuals optimize a forecasting heuristic with a genetic algorithms optimization procedure. We show that this model, taken almost directly from Anufriev et al. (2013), is able to replicate well the main findings of an experimental study by Assenza et al. (2013). Our contribution therefore is that we show that the discussed learning model can be extended from the context of micro to macro economies.

In dynamic macroeconomic models, such as the standard New Keynesian model, expectation feedback plays an important role in the shape and stability of economic equilibria. Traditional literature (Muth, 1961) would disregard the potential heterogeneity of forecasting behavior and focus instead on the model consistent Rational Expectations. However, limitations on the individual rationality are likely to prevail (Cornea et al., 2013), especially when the individuals need time to learn the structure of the economy (Sargent, 1993). Once we allow for bounded rationality, a non-linear price expectations feedback can lead to complicated and potentially volatile dynamics (Anufriev et al., 2011). On the other hand, there are many different forecasting rules that individuals can use to form their expectations about future prices. It is therefore important to study how such rules are selected in a realistic learning environment; and how does the learning in the context of macroeconomics relate to forecasting in other economic settings, including financial or commodity markets.

In order to study the individual forecasting behavior, literature suggests Learning-to-Forecast (LtF) experiments (Marimon et al., 1993). In these, the role of human subjects is to forecast prices, which are hence translated into realized prices through some market mechanism, such as a simple supply-driven economy in which the subjects are framed as advisers to the commodity producers. The LtF experiments typically have a straightforward and unique fundamental equilibrium, and hence can be directly used to assess individual learning dynamics. In practice, they show that individuals indeed have heterogeneous expectations (Anufriev and Hommes, 2012; Heemeijer et al., 2009; Hommes, 2011), which greatly depend on the specific structure of the feedback market. Moreover, subjects can coordinate away from the fundamental equilibrium, or even on oscillatory time paths (Assenza et al., 2013).

In order to understand this heterogeneous behavior, the LtF experimental data can be used to construct and assess learning models. A notable example is work by Anufriev and Hommes (2012), who adapted the Brock-Hommes model (Brock and Hommes, 1997) into a Heuristic Switching Model (HSM) with four simple rules, and apply it with success to the experiment of Hommes et al. (2005). Assenza et al. (2013) use the same model to explain their experimental findings. In general, HSM remains a versatile model that can approximate the individual learning of forecasting behavior across different experiments.

A generalized, agent-based counterpart of HSM is the model of individual learning based on genetic algorithms (GA; Haupt and Haupt, 2004). GA is an optimization method based on a population of arguments which compete on their function value and can therefore be applied to a wide class of problems: they rely on an intelligent search of a large but finite solution space using statistical methods and can deal with discrete variables and noncontinuous cost functions (Haupt and Haupt, 2004). Arifovic (1991) has developed an augmented GA model, which was consequently applied in different economic settings such as a cobweb model (Arifovic, 1994; Hommes and Lux, 2013) and an overlapping generation model (Arifovic, 1995). Following a more mature version of the model by Hommes and Lux (2013), Anufriev et al. (2013) have shown that a model, in which individuals independently optimize their prediction rules using GA, is able to replicate experimental findings from four different LtF experiments, based on commodity or financial markets. A great advantage of this approach is that this model is a generalized version of the HSM with no need a pre-specification of forecasting rules, and hence can be used to motivate the parametrization of the latter model (Anufriev et al., 2013).

This paper follows up on Assenza et al. (2013) by using a GA model to replicate the results from their LtF experiment based on New Keynesian macro model. We use the same GA model as Anufriev et al. (2013), i.e. we update heuristics with an adaptive, an anchor an a trend extrapolation coefficient. In this way we contributed to understanding LtF experiments in a New Keynesian environment using GA. Unlike Arifovic et al. (2012) who have investigated GA in a New Keynesian environment as well, we explain the heterogeneous behavior with heuristics that depend only on the realizations and previous predictions by the agent, similar to the heuristics used by Heemeijer et al. (2009), Anufriev and Hommes (2012) and Assenza et al. (2013). Moreover, we use experimental settings with different types of complexity and show that LtF experiments with increasing complexities can be explained using the same GA model. This shows that the GA model is versatile and can replicate varied experimental economies, but also remains robust against the *object of the forecasting task*. To be specific, our main contribution is to prove that the original GA model by Anufriev et al. (2013) explains the individual learning to forecast not only of prices in a specific market, but also of macro variables such as inflation and output gap. However, we note that the more complicated experimental treatments require some adaptation of the GA model (*c.f.* Anufriev et al., 2013, who encountered a similar problem with the two-period ahead non-linear asset pricing economy). This shows that the further research should focus on extensions of this model.

We replicate results for six different experimental treatments: three different treatments with increasing complexity (1, 2 and 3), each subdivided into two sessions (A and B). The results from the treatments 1, 2 and 3 in the experiment can be classified in three types of aggregate behavior respectively: converging, oscillatory and dampened oscillatory behavior. The main goal of this paper is to show that all three types of behavior can be reproduced using one GA model. We use Monte Carlo simulations in the spirit of Anufriev et al. (2013) to investigate how well the model explains the experimental data.

2 Model

During the past decade the New Keynesian (NK) monetary model has been a widely used framework for the analysis of monetary policy, in which inflation expectations play an important role. Branch and McGough (2009) have incorporated bounded rationality at the individual agent level and heterogeneous expectations in the NK model. Assenza et al. (2013) use this model to set up a laboratory experiment. In order to study the individual expectations process, subjects are asked to forecast the inflation rate under three different scenarios.

This section describes the NK model, the experimental setup and an explanation of the experimental results, followed by a description of the GA model.

2.1 New Keynesian model

The New Keynesian model with heterogeneous expectations developed by Branch and McGough (2009) is described by the following equations:

$$y_t = \overline{y}^e_{t+1} - \varphi(i_t - \overline{\pi}^e_{t+1}) + g_t \tag{1}$$

$$\pi_t = \lambda y_t + \rho \overline{\pi}^e_{t+1} + u_t \tag{2}$$

$$i_t = \overline{\pi} + \phi_\pi(\pi_t - \overline{\pi}) \tag{3}$$

In this system equation 1 describes the aggregate demand in which the output gap y_t depends on the average expected output gap $\overline{y}^e_{t+1}$ and on the real interest rate $i_t - \overline{\pi}^e_{t+1}$. Equation 2 shows how the inflation rate depends on the output gap and on average expected inflation. Equation 3 is the monetary policy rule implemented by the monetary authority in order to keep inflation at its target value $\overline{\pi}$. In equations 1 and 2, g_t and u_t are small normally distributed errors.[1]

The NK model requires agents to forecast both inflation and the output gap. Given that forecasting variables might be a too difficult task for subjects, the experiment is ran using three different treatments.

2.2 Treatments

In the first treatment of the experiment where only the inflation rate needs to be forecast, the model reduces to a framework with a structure similar to the experimental framework that was used by Anufriev et al. (2013). In this treatment, the output gap is fixed at the equilibrium predictor. In the second treatment subjects only forecast the inflation rate and expectations on the output gap are represented by naive expectations. This results in a two dimensional structure which makes the framework more complicated. The third treatment of the experiment represents an economy driven by individual expectations on two different aggregate variables, with two different groups of forecasters, predicting respectively inflation and the output gap.

Moreover, all treatments are run under different monetary policy regimes, a regime *a* in which $\phi_\pi = 1$ and a regime *b* in which $\phi_\pi = 1.5$.

[1] In the experiment the parameters are fixed at the calibration by Clarida et al. (2000): $\rho = 0.99$, $\varphi = 1$ and $\lambda = 0.3$. The inflation target is set to $\overline{\pi} = 2$. Coefficient ϕ_π measures the response of nominal interest rate i_t to deviations of the inflation rate π_t from its target $\overline{\pi}$.

2.2.1 Treatment 1

In the first treatment subjects forecast inflation, while the expectations on the output gap are assumed to be given by the equilibrium predictor $\overline{y}^e_{t+1} = (1-\rho)\overline{\pi}\lambda^{-1}$. The initial set of equations can now be written as:

$$
\begin{aligned}
(4) \qquad y_t &= (1-\rho)\overline{\pi}\lambda^{-1} - \varphi(i_t - \overline{\pi}^e_{t+1}) + g_t \\
(5) \qquad \pi_t &= \lambda y_t + \rho\overline{\pi}^e_{t+1} + u_t \\
(6) \qquad i_t &= \overline{\pi} + \phi_\pi(\pi_t - \overline{\pi})
\end{aligned}
$$

in which $\overline{\pi}^e_{t+1}$ is the average prediction of the subjects in the experiment. Substituting 6 into 4 results in:

$$
\begin{aligned}
(7) \qquad y_t &= (1-\rho)\overline{\pi}\lambda^{-1} + \varphi\overline{\pi}(\phi_\pi - 1) - \varphi\phi_\pi\pi_t + \varphi\overline{\pi}^e_{t+1} + g_t \\
(8) \qquad \pi_t &= \lambda y_t + \rho\overline{\pi}^e_{t+1} + u_t
\end{aligned}
$$

Solving this in terms of inflation π_t gives:

$$
(9) \qquad \pi_t = \frac{(1-\rho) - \lambda\varphi(\phi_\pi - 1)}{1+\lambda\varphi\phi_\pi}\overline{\pi} + \frac{\lambda\varphi + \rho}{1+\lambda\varphi\phi_\pi}\overline{\pi}^e_{t+1} + \frac{\lambda g_t + u_t}{1+\lambda\varphi\phi_\pi}
$$

Note that this is a linear relation between π_t and $\overline{\pi}^e_{t+1}$, plus a composite shock as a third term.

2.2.2 Treatment 2

In the second treatment subjects also forecast inflation, but now the expectations on the output gap are assumed to be represented by naive expectations: $\overline{y}^e_{t+1} = y_{t-1}$. In this case, inflation and output gap at time t can be written as follows:

$$
\begin{aligned}
(10) \qquad y_t &= \varphi\overline{\pi}(\phi_\pi - 1) - \varphi\phi_\pi\pi_t + \varphi\overline{\pi}^e_{t+1} + y_{t-1} + g_t \\
(11) \qquad \pi_t &= \lambda y_t + \rho\overline{\pi}^e_{t+1} + u_t
\end{aligned}
$$

in which $\overline{\pi}^e_{t+1}$ is the average prediction of the subjects in the experiment. In matrix form, this system of equations becomes:

$$
(12) \qquad \begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \frac{1}{1+\lambda\varphi\phi_\pi}\left(\begin{bmatrix} 0 & \varphi(1-\phi_\pi\rho) \\ 0 & \lambda\varphi + \rho \end{bmatrix} \begin{bmatrix} \overline{y}^e_{t+1} \\ \overline{\pi}^e_{t+1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \lambda & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & -\varphi\phi_\pi \\ \lambda & 1 \end{bmatrix} \begin{bmatrix} g_t \\ u_t \end{bmatrix} \right)
$$

This treatment is more complicated than the first since the inflation does not only depend on expected inflation but also on the output gap in the previous period.

2.2.3 Treatment 3

In the third treatment, two groups participate in the same economy, where one group forecasts inflation and the other group forecasts the output gap. Substituting 3 into 1 and writing the equations in matrix form results in the following system:

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \frac{1}{1+\lambda\varphi\phi_\pi} \left(\begin{bmatrix} 1 & \varphi(1-\phi_\pi\rho) \\ \lambda & \lambda\varphi+\rho \end{bmatrix} \begin{bmatrix} \overline{y}^e_{t+1} \\ \overline{\pi}^e_{t+1} \end{bmatrix} + \begin{bmatrix} 1 & -\varphi\phi_\pi \\ \lambda & 1 \end{bmatrix} \begin{bmatrix} g_t \\ u_t \end{bmatrix} \right) \tag{13}$$

In contrast to the first two treatments, treatment 3 represents an experimental economy that is driven by individual expectations on two different interacting aggregate variables.

2.2.4 Monetary policy regimes

Each of the three treatments was run under two different monetary policy regimes. In sessions *a* coefficient ϕ_π is set to 1 so that equation 3 reduces to $i_t = \pi_t$. With this setting, there is no attraction whatsoever to target inflation value $\overline{\pi}$. In session *b* coefficient ϕ_π is set to 1.5 so that the monetary policy responds to inflation aggressively. In this scenario, $\overline{\pi}$ does not drop out of equation 3, so that the inflation rate has a tendency towards target value $\overline{\pi}$.

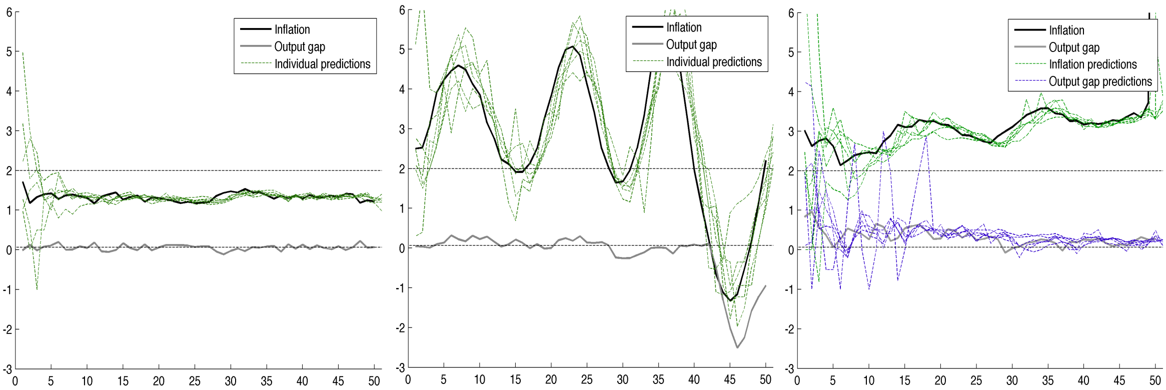
2.3 Experimental results

The experimental results can be classified into three different types of behavior. The experimental results that clearly show these different types of behavior are presented in figures 1a through 1f.

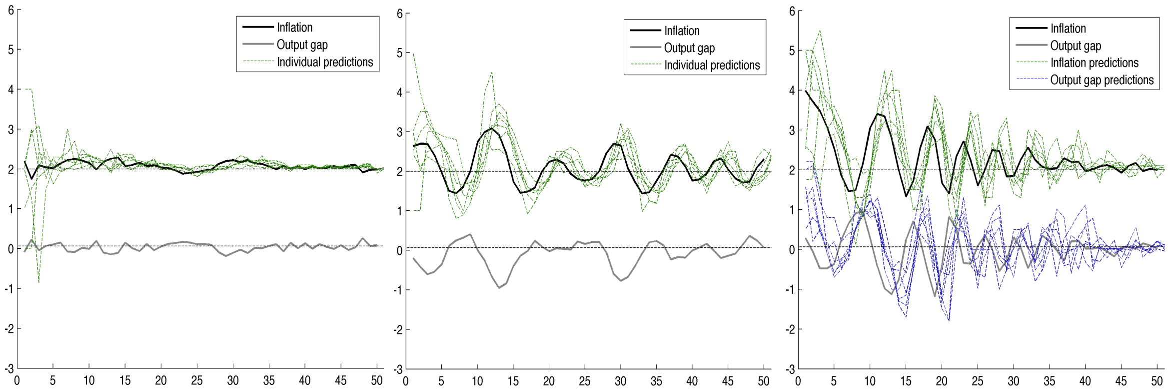
Convergence: In treatment 1a, two groups converge to a non-fundamental steady state equilibrium. Because the monetary policy responds weakly to inflation rate fluctuations, subjects coordinate on inflation rates other than the target inflation rate. In treatment 1b, two out of three group also converge to a steady state. In this case however, the monetary policy responds aggressively to inflation, so that subjects tend to coordinate on the target inflation.

Oscillations: In the second treatment, a different type of aggregate behavior can be observed. Subjects in group 2 of treatment 2a converge to an oscillatory pattern. Towards the end of the session the oscillations are principally above the target inflation rate, due to the monetary policy settings. In treatment 2b, subjects are again forced towards the target inflation rate. The experiment shows small oscillations around the fundamental inflation rate.

Dampened oscillations: The third type of aggregate behavior that this research aims to reproduce is an oscillatory convergence towards a steady state. This behavior occurs in the



(a) Treatment 1a, group 3 **(b)** Treatment 2a, group 2 **(c)** Treatment 3a, group 2



(d) Treatment 1b, group 2 **(e)** Treatment 2b, group 3 **(f)** Treatment 3b, group 2

Figure 1: Typical experimental results.

second session of treatment 3b, which starts out with oscillations around the target inflation rate. The oscillations slowly dampen such that there is convergence to the fundamental steady state (which is in line with the monetary policy settings of the b-treatment) near the end of the session.

2.4 The genetic algorithm model

We follow Arifovic' augmented GA model, in which every individual starts with a set of forecasting heuristics, for either inflation rate or the output gap, which are encoded in binary string. After initialization of the model the heuristics undergo a GA iteration. This iteration, an optimization procedure that uses four evolutionary operators, is the core of the model.

In the GA model, each individual possesses a population of forecasting heuristics, in which one or more parameters need to be optimized. Every heuristic therefore entails a candidate vector of optimization parameters encoded in a binary string. This binary string can be seen as a chromosome containing one or more genes - the parameters in the vector. Parameter $\theta^n_{h,i,t}$ is the n^{th} parameter in heuristic h of individual i in period t, and is coded in a binary string of length l with binary values $g^{n,k}_{h,i,t}$ at the k^{th} position in the string as follows:

$$\theta^n_{h,i,t} = a_n + \frac{b_n - a_n}{2^{l-1}} \sum_{k=1}^{l} g^{n,k}_{h,i,t} 2^k - 1 \tag{14}$$

Since each gene has a finite length, the parameter values are limited to a finite interval, with a_n and b_n as lower and upper boundary respectively, and to a finite number of different values. The size of this interval, together with the length of the string, determines the precision of the parameter.[2]

2.4.1 GA iteration

The encoded heuristic goes through four stages of updating: reproduction, mutation, crossover and election. The operators in the GA iteration are inspired by the theory of evolution, but also have an economic intuition.

The first operator in every GA iteration is the reproduction operator, which randomly draws 20 heuristics for each individual. Every draw takes places according to the heuristics' probabilities to be chosen for reproduction, based on their performance measure The reproduction operator represents the phenomenon that more successful strategies (in terms of utility) are more likely to be used in the future.

After reproduction there is a small probability that a *mutation* will occur in the new strategy. In the binary string, each position has an equally small chance of changing from a 0 to a 1 or vice versa. Depending on the position of the string in which the mutation takes place, the effect of a single mutation can be significant or very small.

Combining two different strategies into new strategies is captured by the crossover operator, whereas the mutation operator models small changes in strategies. All 20 heuristics that are picked in the reproduction stage are, after mutation, signed up as random pairs and will interchange a part of the binary string that represents the forecasting coefficients.[3]

In the crossover and mutation stage two newly formed heuristics are formed from the two old heuristics for the new period. Because these two new heuristics do not always perform better than its predecessors, an election operator tests the performance of the two new and the two old heuristics. The performance of these strings will be based on the difference of the inflation (or output gap) prediction with respect to the last observed inflation (output gap). Out of these four strings, the best performing two will be chosen for the next period.

After the GA iteration, every individual has an updated set of heuristics at their disposal. From this new set of heuristics, every individual chooses one as their inflation (or output gap) forecast. In order to make this choice, individuals make use of a performance measure for every heuristic: heuristics that perform better according to this measure, have a higher probability of being chosen. The forecasts of all individuals together determine the first actual value for inflation or output gap according to equations 1, 2 and 3. Individuals now take their updated set of heuristics to the next period, in which the next GA iteration starts. The complete

[2] We take genes to consists of 20 bits. This means that interval $[a_n \;\; b_n]$ is subdivided into $2^{20} \approx 10^6$ possible values. In an interval with length 1, say an interval $[0 \;\; 1]$, the precision in the GA model is then approximately 10^{-6}.

[3] In this research the heuristic is a binary representation of multiple coefficients. The crossover is designed such that a subset of these coefficients are interchanged, i.e. the string is not broken up within a coefficient, which means that no new values for coefficients come into existence during the crossover stage. Hence, mutation and crossover have strictly different functions.

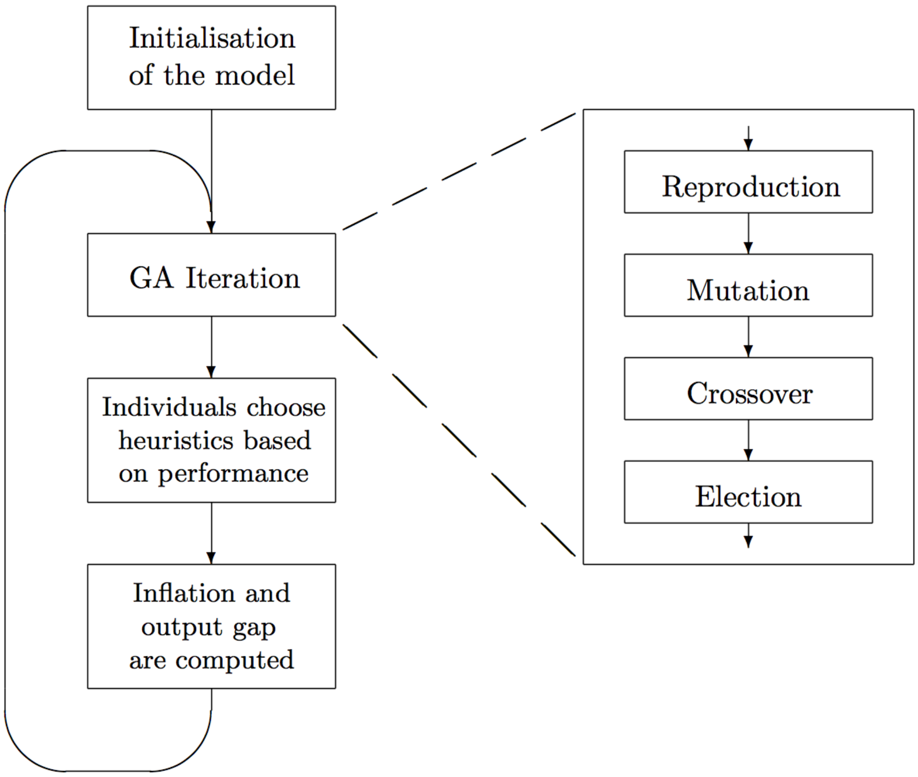


Figure 2: Schematisation of the GA model.

iterative process in the model is illustrated in figure 2. This process continues for a number of periods, in which individuals should be able to make better predictions as time goes by, because their heuristics update every iteration.

2.4.2 Forecasting Heuristics

GA can be interpreted as a generalized version of the forecasting heuristics in the heuristic switching model used by Assenza et al. (2013). Agents do not choose from a set of predefined rules, but use Genetic Algorithms to optimise over a set of coefficients in a simple linear prediction rule, based on the past inflation and/or output gap, individual past prediction, the observed trend and the average of past inflation and/or output gap.

This general rule consists of an adaptive component (α), a trend extrapolating component (β) and an anchor component (γ):

$$x^e_{i,t+1} = \gamma x^{av}_{t-1} + (1-\gamma)(\alpha x_{t-1} + (1-\alpha)x^e_{i,t}) + \beta(x_{t-1} - x_{t-2}) \tag{15}$$

This rule is in line with the so-called 'first order heuristic', which is used by Heemeijer et al. (2009) to explain the participants' behavior in an experimental economy, which is also used in the GA model by Anufriev et al. (2013) to explain behavior in an experimental economy. A condition of the first order heuristic is that the coefficients for the anchor [γ], the last observed value [$(1-\gamma)\alpha$] and the last forecast [$(1-\gamma)(1-\alpha)$] are non-negative and sum up to one. This particular way of formulating the forecasting heuristic ensures this for all values of α and

γ between 0 and 1. Individuals optimize the three coefficients α, β and γ, encoded in a 60-bit string (3×20 bits).

The forecasting heuristics described above are used for simulations of all three different treatments. In treatment 1 and 2, only inflation π is predicted by the participants, so that the forecasting heuristic simply becomes:

$$(16) \qquad \pi^e_{i,t+1} = \gamma_{i,h,t}\pi^{av}_{t-1} + (1-\gamma_{i,h,t})(\alpha_{i,h,t}\pi_{t-1} + (1-\alpha_{i,h,t})\pi^e_{i,t}) + \beta_{i,h,t}(\pi_{t-1} - \pi_{t-2})$$

In treatment 3, however, there are six participants who predict inflation while six other participants predict the output gap. Both variables are updated using the same general rules. Superscripts π and y are added to coefficients α and β to differentiate between inflation and output gap.

$$(17) \qquad \pi^e_{i,t+1} = \gamma^{\pi}_{i,h,t}\pi^{av}_{t-1} + (1-\gamma^{\pi}_{i,h,t})(\alpha^{\pi}_{i,h,t}\pi_{t-1}(1-\alpha^{\pi}_{i,h,t})\pi^e_{i,t})\beta^{\pi}_{i,h,t}(\pi_{t-1} - \pi_{t-2})$$

$$(18) \qquad y^e_{i,t+1} = \gamma^{y}_{i,h,t}\pi^{av}_{t-1} + (1-\gamma^{y}_{i,h,t})(\alpha^{y}_{i,h,t}y_{t-1}(1-\alpha^{y}_{i,h,t})y^e_{i,t})\beta^{y}_{i,h,t}(y_{t-1} - y_{t-2})$$

2.4.3 Performance measure

In this GA framework, every individual has a whole range of forecasting heuristics at hand to forecast inflation (or the output gap) in every period. This choice is made on the basis of the performance of the heuristics, determined by a fitness measure. Hence, the type of performance measure used in the model is of key importance to the simulation process. In the GA model this performance measure is assumed to be equal to the payoff function used by Assenza et al. (2013) in their experiment, namely:

$$(19) \qquad U_{i,h,t} = \frac{100}{1 + \|x^e_{i,h,t-1} - x_{t-1}\|}$$

The performance measure of each heuristic is used twice in every GA iteration: to choose heuristics for reproduction, and to pick one heuristic as a forecast for the next period. In both cases, the probability that a heuristic is chosen is obtained by formalizing the logit-transformation of the utility measure and adding an intensity of choice parameter β_s.[4] This parameter measures the sensitivity of individuals to differences in the performance of their heuristics. This is in line with the performance measure that was used in the HSM by Assenza et al. (2013). The probability that a heuristic is chosen then becomes:

$$(20) \qquad \Pi_{i,h,t} = \frac{\exp(\beta_s U_{i,h,t-1})}{\sum_{h=1}^{H} \exp(\beta_s U_{i,h,t-1})}$$

[4] Subscript s is added to this parameter to distinguish the intensity of choice parameter from the trend extrapolation parameter in the heuristics.

For all simulations in this research this normalized logit-transformation is used. We choose $\beta_s = 1$ in all simulations. The performance measure then simply becomes:

$$\Pi_{i,h,t} = \frac{\exp(U_{i,h,t-1})}{\sum_{h=1}^{H} \exp(U_{i,h,t-1})} \quad (21)$$

2.5 Parametrization

Besides the GA operators, the forecasting heuristics and the performance measure, the model requires the tuning of a few important model settings to replicate the experimental economies. These aspects are discussed in this section.

2.5.1 Initialization of the model

Each session in the LtF experiment consists of 50 periods, which means that the GA model should run for the same amount of time. In order not to use any experimental outcomes, 50-period ahead simulations are carried out. This means that no information from the experiment is used except for the initial predictions by the subjects.[5] The initial predictions in the simulations are chosen equal to the initial predictions from the experiment. After this, the model is ran for 49 more periods, following the structure shown in figure 2.[6]

2.5.2 Parameters GA model

As explained earlier, the parameters α, β and γ are restricted to a finite interval. For α and γ these ranges are obvious: the conditions of the first order rule dictate that both α and γ should be between 0 and 1. The ranges for trend extrapolation parameter β are however not subject to any constrains. Furthermore, trend extrapolation parameters can in theory be negative as well, indicating 'contrarian' behavior. Massaro (2012) showed that subjects in the experiment indeed make use of trend extrapolation, but that virtually all subjects use positive coefficients. The range of β is therefore set to $[0, 3]$ to allow for relatively weak (0 to 1) and relatively strong (1 to 3) trend extrapolation.

Furthermore, the mutation and crossover operator are both subject to a certain probability of occurrence. Throughout the simulations in this research, the mutation rate is set to 0.01, so that during every GA iteration, every bit in every string has a 1% chance of mutating.

[5]In this economic setup, subjects do not predict the present inflation, but the inflation for the next period (i.e. two periods ahead forecasts). Therefore, predictions are needed not only for the first but also for second period. In the second period, subjects observe the inflation in period 1, which they can use in their prediction of the inflation in period 3.

[6]The first period of the iteration, however, differs from the following 48 periods, because subjects cannot use trend extrapolation yet. In this period there is only one observation on the inflation rate. It is therefore assumed that the trend extrapolation coefficient equals zero in this period. After period two, the GA updating continues in full form, where agents can forecast using both past inflation rate and the past trend in the inflation rate.

Crossover does not always happen either; the crossover rate is set to 0.9, which means that each pair of strings has a 90% chance of interchanging a part of the string.

2.5.3 Monte Carlo simulations

In order to investigate how well the model explains the experimental data, Monte Carlo simulations are carried out. For all six treatments, a thousand replications are run, each with the same initial conditions. Monte Carlo simulations enable the composition of confidence intervals. The next section shows the experimental results compared to the mean, a 90% and a 95% confidence interval. Moreover, a confidence interval for some descriptive statistics can be created. For each treatment this enables us to compare the mean, standard deviation and autocorrelation of all thousand replications to the experimental result.

3 Results

Overall, the model is able to replicate three general types of behavior from the experimental results. This section compares the experimental results to the Monte Carlo simulations using graphically constructed confidence intervals of the simulations. Additionally, confidence intervals of the mean, standard deviation and autocorrelation are compared.

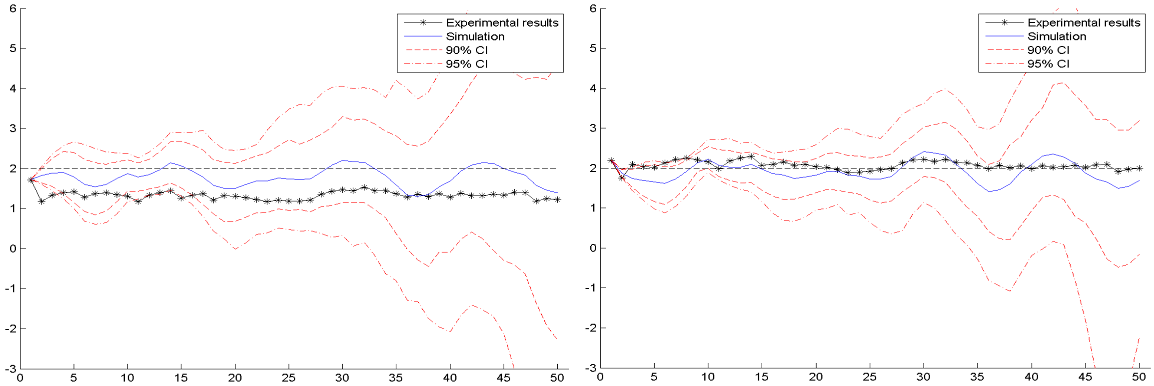
3.1 Treatment 1

In treatment 1, the experiment shows how the subjects coordinate on a steady state. In treatment 1a the monetary policy settings are such that there is no tendency towards the target inflation value, so that coordination on a wide range of inflation rates can occur. In treatment 1b however, the target inflation plays a role in the system of equations, so that subjects coordinate on this value when this converging behavior takes place. The results in Figure 3b clearly show a smaller distribution indicating that GA agents are also drawn towards the target inflation rate.

We also see that the mean of the treatment 1b simulation (Figure 3b) moves around the target inflation, whereas the mean stays below the target inflation in treatment 1a (Figure 3a). In both treatment 1a and 1b, there is however some oscillatory behavior visible, whereas this behavior is absent in the experimental result. Moreover, we see that the standard deviation of the Monte Carlo simulations increases at the end of the simulation period. This indicates at least in some simulations diverging behavior.

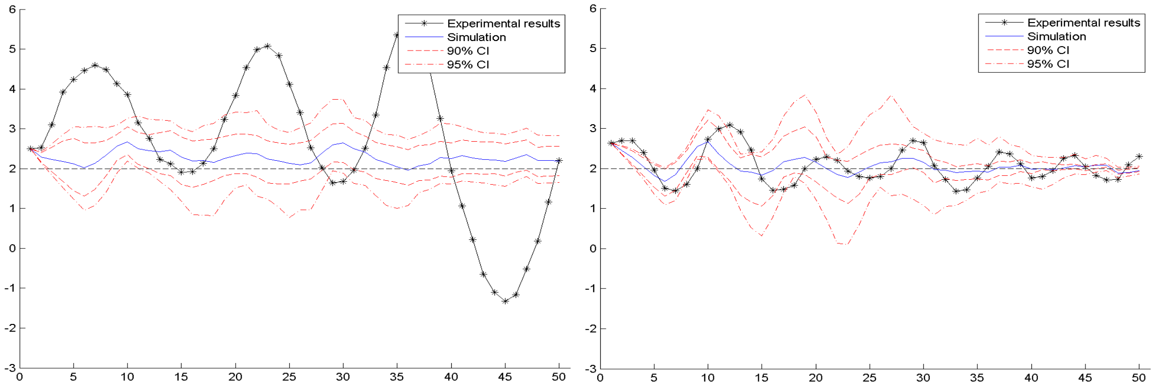
3.2 Treatment 2

The oscillatory behaviour in treatment 2 is captured well in treatment 2b (Figure 4b). In treatment 2 the oscillatory pattern of the Monte Carlo mean matches the experimental results



(a) Treatment 1a **(b)** Treatment 1b

Figure 3: Monte Carlo results: treatment 1.



(a) Treatment 2a **(b)** Treatment 2b

Figure 4: Monte Carlo results: treatment 2.

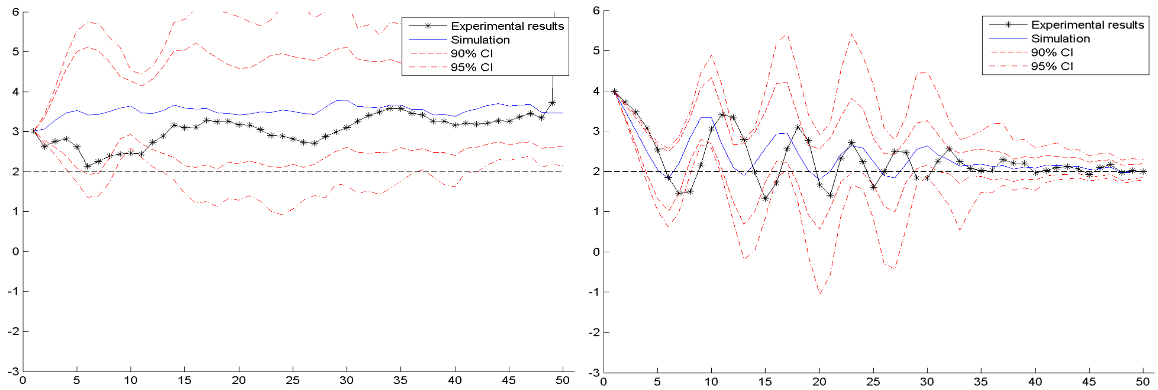
well. Although the amplitude of the mean is smaller (perhaps due to interference), the experimental result remains almost entirely within the 95% confidence interval. Also, just like in treatment 1, the difference in monetary policy settings is captured well.

The large oscillations of treatment 2a (Figure 4a) can however not be replicated by the this model.[7] Secondly, the oscillations dampen towards the end of the simulation period. This is especially visible by the confidence intervals in treatment 2b. This dampening behavior does not clearly occur in the experiments.

3.3 Treatment 3

The experimental results of treatment 3a clearly shows no attraction to the target inflation (Figure 5a). This is replicated well by the simulations. The simulations of treatment 3b replicate the oscillatory behavior that occurs in the experiment (Figure 5b). Similar to treatment 2b in Figure 4b, we see that almost the experimental inflation remains almost entirely within the 95% confidence interval, while the mean of the Monte Carlo simulations closely resembles the experiment. We also see that the dampening of the oscillations resembles the dampening in the experiment.

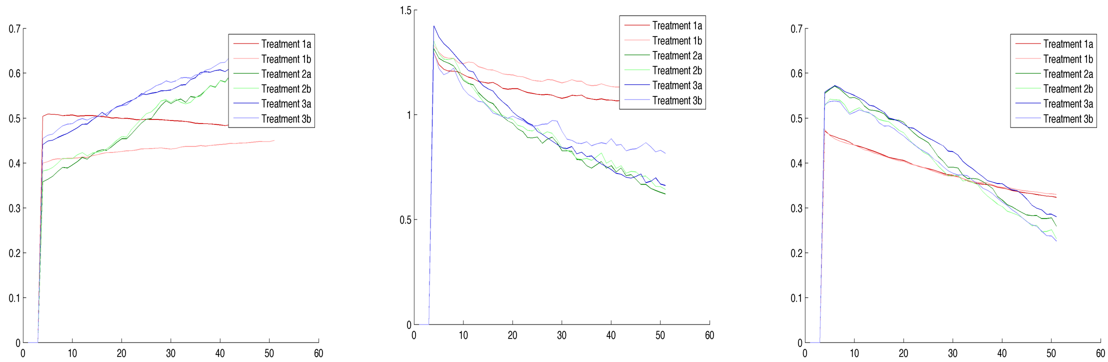
[7]That is, using these specific parameters for the ranges of the optimization parameters, mutation rate, crossover rate etc.



(a) Treatment 3a

(b) Treatment 3b

Figure 5: Monte Carlo results: treatment 3.



(a) Adaption parameter α

(b) Trend extrapolation parameter β

(c) Anchor parameter γ

Figure 6: Monte Carlo results: evolution of the chosen heuristic coefficients.

3.4 Heuristic parameters chosen by the agents

Because each treatment yields different dynamics, coefficients α, β and γ are optimized differently by the agents. Figures 6a through 6c compares the mean α, β and γ in the six different treatments. The x-axis shows the simulation period. In general we can say that each period in the simulation, the adaption parameter α, becomes more important while trend extrapolation β and anchor γ become less important. This goes for all treatments. We do however see differences between the different treatments. What strikes especially is that for each of the coefficients treatment 1 differs from treatments 2 and 3. The steeper curves of the optimization parameters in treatments 2 and 3 indicates that the model incentivizes GA agents to update their coefficients faster in these treatments.

3.5 Descriptive statistics

In addition to the simulation confidence intervals we make a comparison of the mean, variance and autocorrelation of the experimental results to and the Monte Carlo simulations. We can use the Monte Carlo simulations to create a mean and a confidence interval for these three statistics. Table 1 shows that the mean of the experimental inflation rate is close to the mean of the simulated inflation rate. The model especially captures the difference between the a- and b-session of the treatments. In the b-sessions, the mean is closer to target inflation rate due to a strict monetary policy. This also causes the smaller confidence intervals.

Table 1: Mean of the simulated and experimental inflation rate

Treatment	1a	1b	2a	2b	3a	3b
Experimental μ	1.3333	2.0719	2.7367	2.0882	3.3728	2.2939
Mean Simulated μ	1.7663	1.9082	2.2653	2.0738	3.5199	2.3500
95% confidence interval	0.8996 – 2.6506	1.3841 – 2.3081	1.6287 – 2.9014	1.9626 – 2.1479	2.1906 – 4.8938	2.1426 – 2.5496

Table 2 shows how the variance of the experimental inflation rate is replicated by the Monte Carlo simulations. We notice that the mean of the variance of the simulations differs from the variance in the experiment and that the confidence intervals of the Monte Carlo simulations are large. In treatment 1 the large confidence intervals are especially striking since the experiments show converging behavior with little variance. In treatments 2 and 3 the mean variance in the simulations is in of same order of magnitude as the variance in the experiment. We see that the model overestimates the variance of the inflation rate in the b-sessions, while it underestimates the variance in the a-sessions.

Table 2: Variance of the simulated and experimental inflation rate

Treatment	1a	1b	2a	2b	3a	3b
Experimental σ^2	0.0111	0.0120	3.2891	0.1937	5.6214	0.3574
Mean Simulated σ^2	3.6695	2.5625	0.9720	0.9818	1.8460	1.5828
95% confidence interval	0.0174 – 30.5718	0.0114 – 25.0909	0.0127 – 10.6451	0.0224 – 12.6180	0.0166 – 18.0184	0.1398 – 13.5817

There is a certain consistency among treatments in the Monte Carlo simulations: the mean simulated autocorrellations of all six treatments are all very close together (with less than 0.1 difference between the highest and lowest value), whereas the experimental autocorrellation varies more. The model's high, consistent autocorrellations suggest that the model entails some oscillatory behavior regardless of the different treatments. The high autocorrellations in treatment 1 with respect to the experimental results are consistent with the oscillations in figures 3a and 3b. Simulations of treatment 2 and 3 perform better, although, the b-session of these treatments appears to be replicated better than the a-session. These are the treatments with the most evident oscillatory behavior.

Table 3: Autocorrellation of the simulated and experimental inflation rate

Treatment	1a	1b	2a	2b	3a	3b
Experimental ρ_1	0.3192	0.4698	0.9404	0.7716	0.4012	0.7047
Mean Simulated ρ_1	0.8234	0.7779	0.7461	0.7368	0.7706	0.7863
95% confidence interval	0.6588 – 0.9355	0.5565 – 0.9157	0.6084 – 0.8654	0.6274 – 0.8282	0.6172 – 0.8919	0.5858 – 0.9601

4 Conclusions and Recommendations

In this paper we present a genetic algorithms model in which individuals optimize an adaptive, a trend following and an anchor coefficient in a population of general prediction heuristics. With this model, based on Anufriev et al. (2013), we replicate results of a Learning-to-Forecast experiment by Assenza et al. (2013). The experiment investigates how the individuals learn to forecast in a New Keynesian economy with three different treatments, each of them with two different monetary policy settings. The results of this experiment can be classified in three types of aggregate behavior: converging, oscillatory and dampened oscillatory behavior.

The model is able to replicate these types of behavior from the experimental results. In the first treatment, which typically shows converging behavior, the model clearly captures the difference between the two monetary policy settings. During the procedure of replicating the oscillatory and dampened oscillatory behavior, we found that the model is sensitive to changes in the allowed ranges for the trend extrapolation coefficient. We use a wider range for this parameter than Anufriev et al. (2013). This underlines their finding that a different feedback structures lead to different degrees of trend extrapolating behavior. As a result, further work may be necessary to find a generalized version of this model that would capture this aspect of the individual behavior.

We show that a single model with a simple set of rules can explain adaptive behavior of human subjects to a predictions feedback with varying levels of complexity. Like Arifovic et al. (2012) this paper contributes to understanding learning behavior in a New-Keynesian environment. It furthermore shows that heterogeneous behavior can be explained by an adaptive, anchor and trend extrapolating component. We also contribute to the existing literature that GA can be used to explain heterogeneous behavior in LtF experiments with different types of complexity, but also for macro experiments next to more classical ones based on commodity and financial markets.

References

Anufriev, M., Assenza, T., Hommes, C., and Massaro, D. (2011). Interest rate rules with heterogeneous expectations. University of Amsterdam.

Anufriev, M. and Hommes, C. (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64.

Anufriev, M., Hommes, C., and Makarewicz, T. (2013). Learning-to-forecast with genetic algorithms. Working paper.

Arifovic, J. (1991). *Learning by genetic algorithms in economic environments*. PhD thesis, University of Chicago, Department of Economics.

Arifovic, J. (1994). Genetic algorithm learning and the cobweb model. *Journal of Economic dynamics and Control*, 18(1):3–28.

Arifovic, J. (1995). Genetic algorithms and inflationary economies. *Journal of Monetary Economics*, 36(1):219 – 243.

Arifovic, J., Bullard, J., and Kostyshyna, O. (2012). Social learning and monetary policy rules. *The Economic Journal*.

Assenza, T., Heemeijer, P., Hommes, C., and Massaro, D. (2013). Individual expectations and aggregate macro behavior. Tinbergen Institute Discussion Paper.

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. and Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65(5):1059–1095.

Clarida, R., Gali, J., and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. *The Quarterly Journal of Economics*, 115(1):147–180.

Cornea, A., Hommes, C., and Massaro, D. (2013). Behavioral heterogeneity in us inflation dynamics. Technical report, Tinbergen Institute Discussion Paper.

Haupt, R. and Haupt, S. (2004). *Practical Genetic Algorithms*. John Wiley & Sons, Inc., New Jersey, 2nd edition.

Heemeijer, P., Hommes, C., Sonnemans, J., and Tuinstra, J. (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic Dynamics and Control*, 33(5):1052–1072.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control*, 35(1):1–24.

Hommes, C. and Lux, T. (2013). Individual expectations and aggregate behavior in learning to forecast experiments. *Macroeconomic Dynamics*, 17(2):373–401.

Hommes, C., Sonnemans, J., Tuinstra, J., and Velden, H. v. d. (2005). Coordination of expectations in asset pricing experiments. *The Review of Financial Studies*, 18(3):pp. 955–980.

Marimon, R., Spear, S. E., and Sunder, S. (1993). Expectationally driven market volatility: an experimental study. *Journal of Economic Theory*, 61(1):74–103.

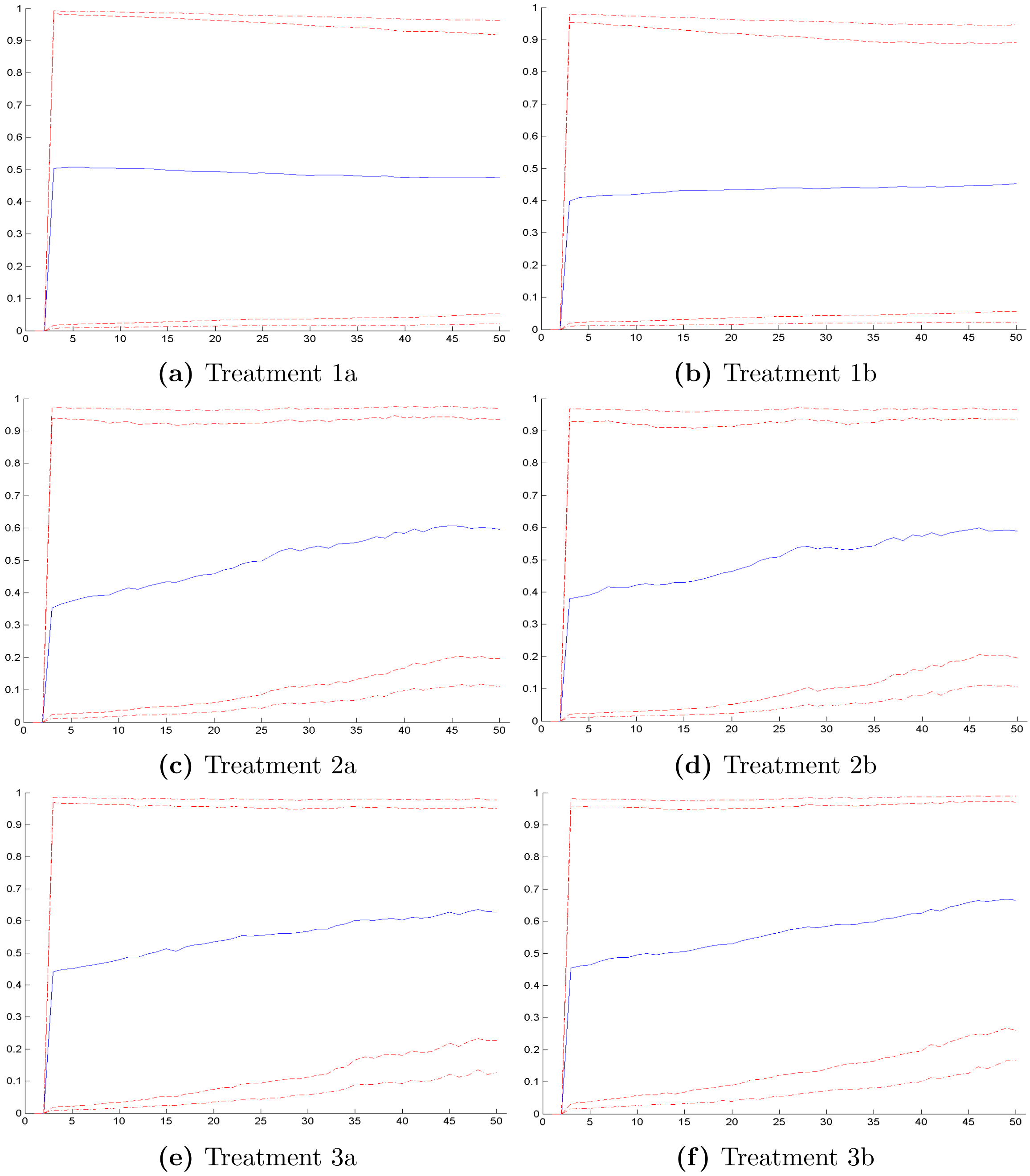
Massaro, D. (2012). *Bounded rationality and heterogeneous expectations in macroeconomics*. Thela Thesis.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29(3).

Sargent, T. J. (1993). Bounded rationality in macroeconomics: The Arne Ryde memorial lectures. *OUP Catalogue*.

Appendices

A Confidence intervals



(a) Treatment 1a

(b) Treatment 1b

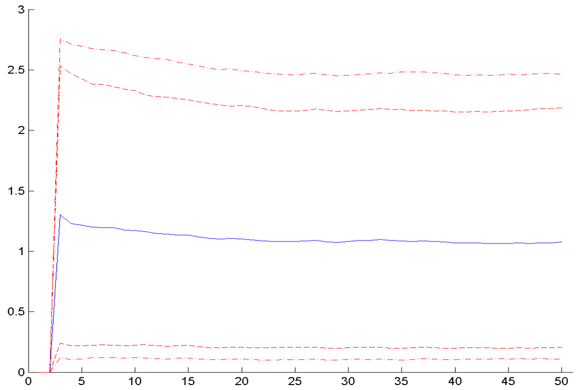
(c) Treatment 2a

(d) Treatment 2b

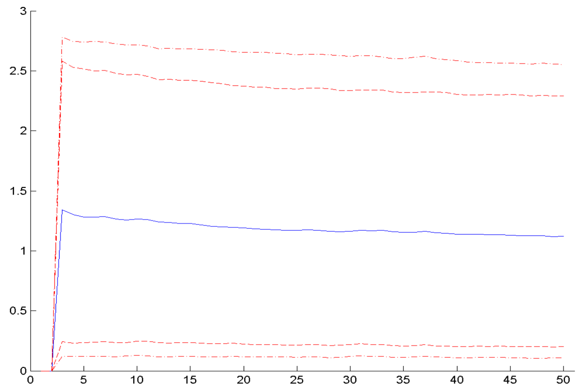
(e) Treatment 3a

(f) Treatment 3b

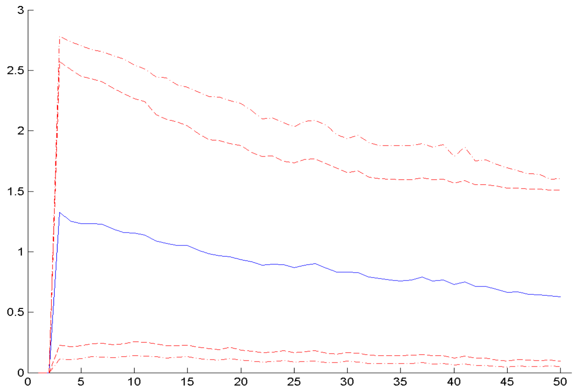
Figure 7: GA model: confidence intervals for the α parameter.



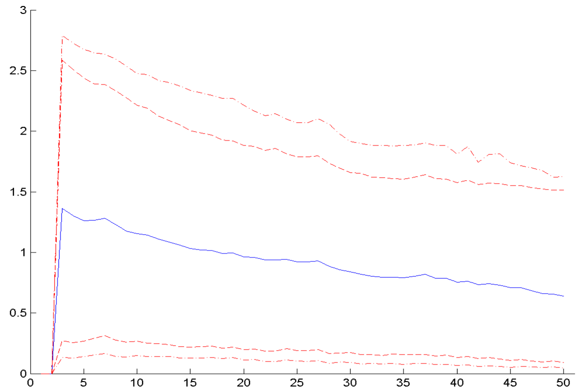
(a) Treatment 1a



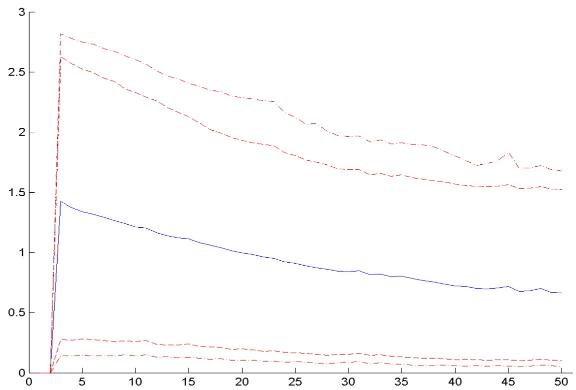
(b) Treatment 1b



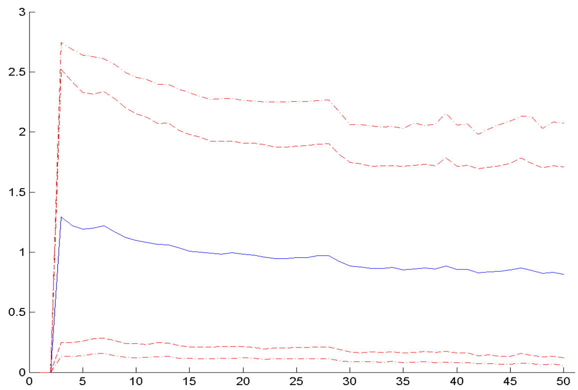
(c) Treatment 2a



(d) Treatment 2b

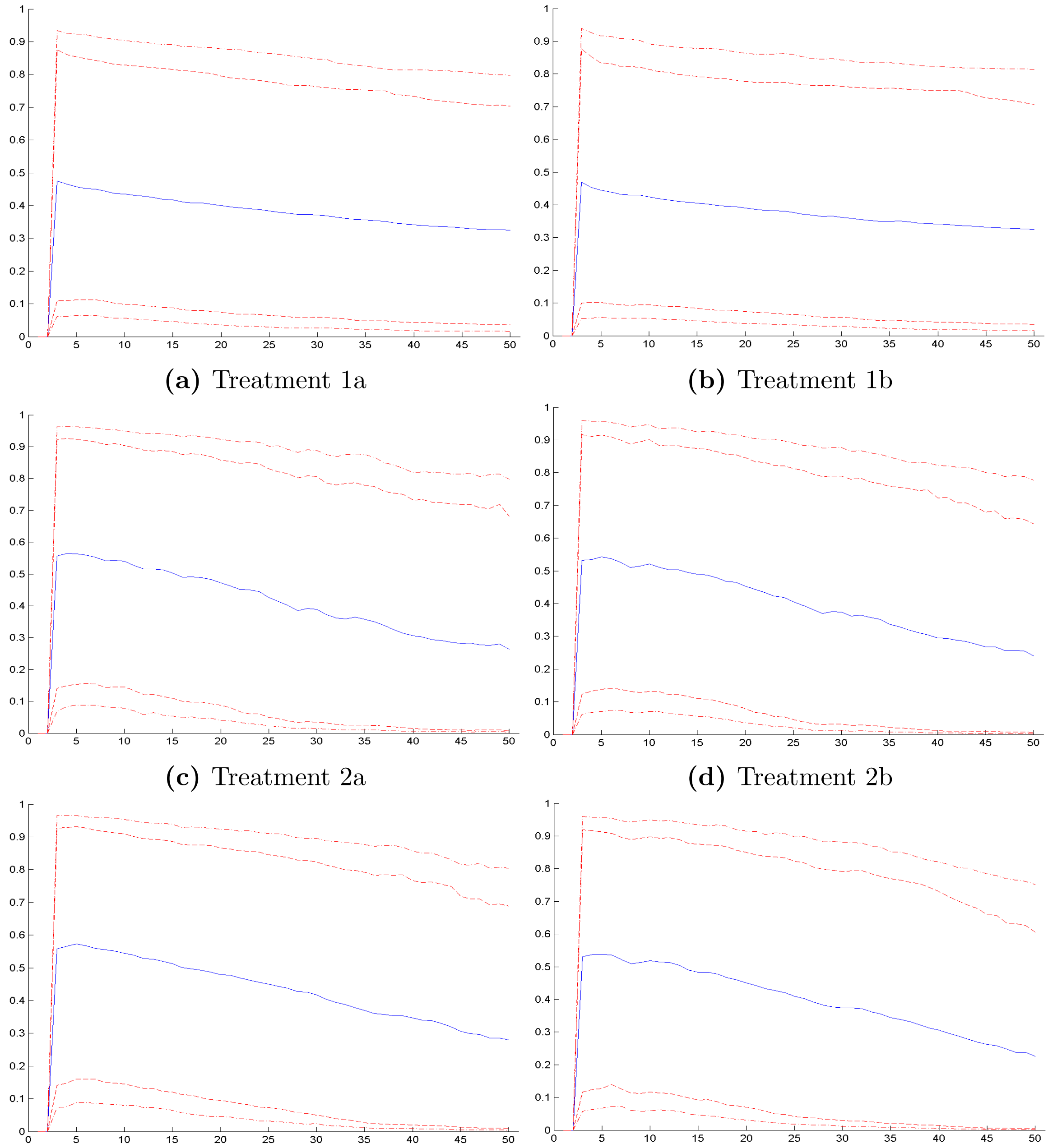


(e) Treatment 3a



(f) Treatment 3b

Figure 8: GA model: confidence intervals for the β parameter.



(a) Treatment 1a **(b)** Treatment 1b

(c) Treatment 2a **(d)** Treatment 2b

(e) Treatment 3a **(f)** Treatment 3b

Figure 9: GA model: confidence intervals for the γ parameter.