Optimal Cheating in Monetary Policy with Individual Evolutionary Learning

Jasmina Arifovic and Olena Kostyshyna*

May 31, 2005

Preliminary and Incomplete
Do Not Quote or Reference or Distribute

*Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby B.C. V5A 1S6, Canada, phone: (604) 291 3508, fax: (604) 291 - 5944, e-mail: arifovic@sfu.ca; okostysh@sfu.ca.

Abstract

We study individual evolutionary learning in the setup developed by Deissenberg and Gonzalez (2002). They study a version of the Kydland-Prescott model in which in each time period monetary authority optimizes weighted payoff function (with selfishness parameter as a weight on its own and agent's payoffs) with respect to inflation announcement, actual inflation and the selfishness parameter. In addition, in each time period agent makes probabilistic decision on whether to believe in monetary authority's announcement. The probability of how trustful the agent should be is updated using reinforcement learning. The inflation announcement is always different from the actual inflation, and the private agent chooses to believe in the announcement if the monetary authority is selfish at levels tolerable to the agent. As a result, both the agent and the monetary authority are better off in this model of optimal cheating.

In our model, both the agent and the monetary authority adapt using a model of individual evolutionary learning (Arifovic and Ledyard, 2004): the agent learns about her probabilistic decision, and the monetary authority learns about what level of announcement to use and how selfish to be. We conduct simulations with two different payoff functions - simple (selfishness weighted payoff from Deissenberg/Gonzales model) and "expected" (selfishness weighted payoffs in 'believe' and 'not believe' outcomes weighted by the probability of agent to believe). The results for the first type of simulations include those with very altruistic monetary authority and the agent that believes the monetary authority when it sets announcement of inflation at sufficiently low levels. In the simulations with "expected" payoffs, monetary authority learned to set the announcement at zero which, in turn, resulted in the zero actual inflation. This Ramsey outcome gives the highest possible payoff to both the agent and the monetary authority. Both types of simulations can also explain changes in average inflation over longer time horizons. When monetary authority starts experimenting with its announcement or selfishness, the agent might change her 'believe' ('not believe') action into the opposite one that entails changes in the actual inflation.

Keywords: monetary policy, optimal cheating, evolutionary learning

Introduction

The problem of the central bank credibility and its ability to commit to policies has long been studied since Kydland and Prescott (1977). Many different extensions and solutions were suggested to deal with the time inconsistency problem and inflation bias arising from it. Barro and Gordon (1983) introduced reputation in the repeated game setup and showed that central bank may deliver low inflation if it needs to establish reputation for delivering low inflation. Backus and Driffill(1985), Ball (1995), Cukierman and Liviatan (1991) developed models with asymmetric information about central bank's information or its preferences (type). Zero-inflation policy is not completely credible when central bank cares about output/unemployment. Similarly, zero-inflation announcement is not credible for agent facing uncertainty about the type of the government (optimizer or inflation-fighter) because both types would want to announce zero-inflation: the first one will do this to get unemployment reduction, the second - because it wants to fight inflation. This announcement does not communicate any information about future government's action, and so agent has no reason to take it into consideration when forming inflation forecast. Backus and Driffill (1985) show that central bank without commitment technology wants to mimic that with ability to commit to low inflation in order to build reputation and then gain by delivering high inflation that reduces unemployment. Cukierman and Liviatan (1991) develop model with announcement of inflation and uncertainty of the public as to which type (optimizer or inflation fighter) central bank is. As long as public assigns positive probability to the central bank being optimizer willing to tolerate higher inflation in order to reduce unemployment, then even inflation-fighter central bank is shown to deliver positive inflation. The other directions of solving time inconsistency problem include assigning and giving independence to the central banker with much stronger than normal preference over low inflation or designing contracts and institutional structure so that the central banker will deliver low inflation. Yet another way to reduce inflation bias arising from discretion and lack of credibility is to limit flexibility of the central bank, i.e. fixed exchange rate or inflation targeting. Inflation targeting is recently used in Canada, Sweden, Finland, the UK, New Zealand.

In the above mentioned literature, it is established that announcements alone are not credible, and only actions are informative: government reveals information about itself by actions. So it adopts lower inflation policy for some time period just to establish reputation and influence agent's beliefs about its type. Announcements can communicate information only if different types of government want to announce different policies. But unlike in separating equilibrium, private agent is uncertain about the government's action after announcement. Stein (1989) establishes that government is better off by making imprecise announcements that are cheap talk. In this way, it gives up its ability to manipulate expectations perfectly but only in some very crude way, and so it becomes able to communicate some information about inflation more credibly.

The role of cheap talk announcements is explored in Deissenberg and Gonzales (2002). They demonstrate that even misleading announcements can be Pareto-improving upon Nash equilibrium solution of Kydland-Prescott model. Announcements are believed because they provide private agent with information about future actions of government and increase welfare of both agent and government and so it becomes costly not to take announcements into account. Deissenberg and Gonzales model government as optimizer, and private agent - as boundedly rational using reinforcement learning. They find that private agent learns to play equilibrium where announcement is not respected, but this allows both agent and government to achieve Pareto-improving result. We explore the properties of this model when both private agent and central bank update their believes using individual evolutionary learning. Our aim is to check whether agent can learn to be cheated upon in order to achieve higher payoff.

The paper is organized in the following way. First, we overview time inconsistency problem and model of Deissenberg and Gonzales. Second, we describe the setup of our simulations. Third, we summarize the regularities of the observations in the simulations and provide explanations of the results.

1 Model.

1.1 Time inconsistency in Kydland-Prescott model.

Kydland and Prescott (1977) developed model with government and private agents that maximize their utilities by choosing inflation and forecast of inflation respectively. Government (L) maximizes utility with respect to actual inflation y:

$$J^L = -1/2(U^2 + y^2) = -1/2((U^* - \theta(y-x))^2 + y^2), \tag{1}$$

where U^* is natural rate of unemployment.
Agents (F_i) maximizes utility with respect to inflation forecast x_i:

$$J_i^F = -1/2((y - x_i)^2 + y^2) = -1/2((y - x_i)^2 + y^2) \tag{2}$$

The agents are symmetric and atomistic in this model, and their forecasts are the same $x_i = x$. The reaction functions are:

$$\begin{aligned} T^L &= argmax J^L = \theta/(1-\theta)(U^* + x), & (3) \\ T^F &= argmax J^F = y & (4) \end{aligned}$$

Agent always makes perfect prediction of government's choice of inflation, this comes from the definition of agent's utility function.
Two solutions were studied: Nash and Stackelberg (L as a leader). In Nash equilibrium, government and agent choose best response to each other's action.

$$\begin{aligned} x^N &= y^N = \theta U^* & (5) \\ J^{L,N} &= J^L(y^N, x^N) = -(1+\theta^2)U^{*2} & (6) \\ J^{F,N} &= J^F(y^N, x^N) = -1/2\theta^2 U^{*2} & (7) \end{aligned}$$

Stackelberg (Ramsey)equilibrium is the outcome of game where government chooses its action knowing how agent will react to this choice with its best response $T^F(y)$.

$$x^{SL} = y^{SL} = 0 \tag{8}$$
$$J^{L,SL} = -1/2U^{*2} \tag{9}$$
$$J^{F,SL} = 0 \tag{10}$$

Both players are better off in Stackelberg equilibrium. But it is time inconsistent: once agent has forecast of zero, government is always better off by setting actual inflation according to its reaction function. Agent know this and so Pareto-inferior Nash equilibrium is the outcome. Kydland-Prescott model is static, and in dynamic setting, Ramsey outcome can be supported by need to have good reputation, incentive contracts or delegation.
In Kydland-Prescott model, agent chooses his expectation x after government made its decision about actual inflation y that makes it necessary to distinguish between announcement of inflation and realized inflation.

1.2 Optimal cheating in Deissenberg-Gonzales model.

Deissenberg and Gonzales (2002) introduce a separate variable called announcement of inflation. The structure of their model includes the following steps:

1. Government makes announcement of inflation, y^a, it is cheap talk.
2. Agent hears this announcement and formes expectation of inflation x.
3. Given agent's expectation of inflation, government chooses actual inflation, y.

So government has 2 instruments: announcement of inflation y^a, and actual inflation y. Announcement influences agent's expectation x. Agent has 2 options: she can believe or discard government's announcement:

- Agent does not believe government, so announcement does not influence expectation at all. The game is Stackelberg with agent as a leader. This outcome will be referred to as SF, or PAL (private agent leadership).

$$x^{SF} = \frac{\theta(1-\theta^2)U^*}{1+\theta^4} \tag{11}$$
$$y^{SF} = \frac{\theta U^*}{1+\theta^4} \tag{12}$$
$$J^{L,SF} = -\frac{(1+\theta^2)U^{*2}}{2(1+\theta^4)^2} \tag{13}$$
$$J^{F,SF} = -\frac{\theta^2 U^{*2}}{2(1+\theta^4)} \tag{14}$$

- Agent believes that government will set actual inflation as announced: $y = y^a$. Agent's reaction function is $T^F(y^a)$. So in step 3, government maximizes with respect to y^a and y with x replaced by $T^F(y^a)$. The outcome is optimal cheating (OC):

$$y^{a,OC} = x^{OC} = -\frac{U^*}{\theta} \tag{15}$$
$$y^{OC} = 0 \tag{16}$$
$$J^{L,OC} = 0 \tag{17}$$
$$J^{F,OC} = -\frac{U^{*2}}{2\theta} \tag{18}$$

Although OC is very attractive for government: $J^{L,OC} = 0$ is the highest payoff government can ever get, it cannot be supported in repeated game because $J^{F,SF} > J^{F,OC}$. When agent does not believe, her utility is higher. So the outcome of the game will be strategies that can bring higher payoffs than Stackelberg with agent as a leader.
The solutions to all these models are represented in Figure 1:

- Big dots represent maximum possible payoffs for government and agent.
- Stackelberg equilibrium SL (SF) is the tangency point of T^F (T^L) with indifference curve of F(L).
- OC is unconstrained optimum for government (best point of its own reaction function).
- Part Γ of contract curve that lies within the lens delimited by F's and L's indifference curves represent efficient and Pareto-improving points (they Pareto-dominate SF).

Characterization of set Γ.
Define function J^α that is the convex combination of J^L and J^F, where α is government's selfishness parameter. It describes to which extent government wants to take into consideration the interests of private agent. If α is equal to 1, government does not take into consideration agent's payoff and is absolutely selfish; if α is equal 0, then government is absolutely altruistic.

$$J^\alpha = \alpha J^L(x,y) + (1-\alpha)J^F(x,y), x = T^F(y^a) \tag{19}$$

Set O is the set of optimal cheating solutions OC^α that is generated when government bases its decisions on J^α, not J^L:

$$O\{(y^{a,\alpha}, x^\alpha, y^\alpha) : (y^{a,\alpha}, y^\alpha) = argmax J^\alpha, \alpha[0,1], x^\alpha = T^F(y^{a,\alpha})\} \tag{20}$$

If $\alpha = 1$, OC^α is the same as OC solution described above; if $\alpha = 0$, solution coincides with unconstrained optimum for F, i.e. origin on the graph. Thus, set O is part of contract curve between OC solution and unconstrained optimum for F.
Set Γ is defined as:

$$\Gamma = (y^{a,\alpha}, x^\alpha, y^\alpha), \alpha[\alpha_1, \alpha_2], \tag{21}$$

where α_1 is value for which government is indifferent between playing SL and OC^{α}, i.e. $J^{L,\alpha} = J^L(x^{\alpha}, y^{\alpha}) = J^{L,SL}$; α_2 is value for which F is indifferent between playing SL and OC^{α}.

$$\alpha_1 = max[0, \frac{1+\theta^4-\sqrt{1+\theta^2}}{1+\theta^4+(\theta^2)\sqrt{1+\theta^2}}] \tag{22}$$

$$\alpha_2 = \frac{1}{1-\theta^2+\sqrt{1+\theta^4}} \tag{23}$$

Cheating is necessary to achieve Pareto-improving outcome. Outcome SF^{α} can be obtained from SF game by imposing that announcement be respected $y^{a,\alpha} = y^a$, and it is the best point of T^F for government maximizing J^{α}. If $\alpha < 1$, F gets better outcome in SF^{α} than in SF. But SF^{α} does not Pareto-dominate SF for $\alpha \neq 0$. And so government has no incentive to respect any benevolent ($\alpha < 1$) announcement or to make announcement knowing that it will have to respect it. The incentive exists if there is possibility of subsequent cheating. And agent has incentive to be cheated upon, because then it gets higher payoff in set Γ. Cheating by itself does not affect players' payoff, as it does not include discrepancy of actual inflation from announcement; it affects agent's payoff through affecting agent's forecast.

Being gullible in this game is very costly for agent, because once he believes government, the latter has incentive to respond with $T^L(x^{\alpha})$, not $T^{L,\alpha}(x^{\alpha})$, and this can hurt agent a lot. Deissenberg and Gonzales run simulations to determine the value of α that can deliver optimal cheating solution in this model. The structure of their simulations is as follows.

1. L chooses $\alpha[0, 1]$ and $y^{a,\alpha}$.
2. F plays OC (optimal cheating accommodate) with probability π and PAL (play as a leader) with probability $1 - \pi$.
3. L observes action of F: if F played OC, i.e $x = x^{\alpha}$, L plays y^{α}; if F something else, L plays $T^L(x)$, and this gives SF payoffs.
4. F revises probability π to play OC based on whether the realized payoff is higher or lower than J^{SF}.

Probability π is updated with reinforcement learning rule that increases the probability to play strategy if agent had good experience (in terms of payoff) with it.

Government sets α by maximizing expected payoff over horizon of 2 periods subject to agent's reinforcement learning rules:

$$\sum_{\tau=t}^{t+1} E(J^L_{\tau}) \rightarrow \max_{\alpha_{\tau}} \tag{24}$$

$$s.t. \pi_{t+1} = \phi(\pi_t, \rho_t, \delta_t) \tag{25}$$

$$
\begin{aligned}
\rho_t &= (max J^F_\tau) - J^F_t, \tau < t, & (26)\\
ifOCwasplayedatt: & & (27)\\
\pi_{t+1} - \pi_t &= -\pi_t \frac{\rho_t}{1+|\rho_t|} if \rho_t \geq 0, & (28)\\
\pi_{t+1} - \pi_t &= -(1-\pi_t)\frac{\rho_t}{1+|\rho_t|} if \rho_t < 0 & (29)\\
ifPALwasplayedatt: & & (30)\\
\pi_{t+1} - \pi_t &= (1-\pi_t)\frac{\rho_t}{1+|\rho_t|} if \rho_t \geq 0 & (31)\\
\pi_{t+1} - \pi_t &= \pi_t \frac{\rho_t}{1+|\rho_t|} if \rho_t < 0 & (32)\\
& & (33)
\end{aligned}
$$

$$
\max_{\alpha_t, \alpha_{t+1}} \sum_{\tau=t}^{t+1} E(J^L_\tau) =
$$

$$
\max_{\alpha_t, \alpha_{t+1}} EJ^L_t + EJ^L_{t+1} =
$$

$$
= \max_{\alpha_t, \alpha_{t+1}} \pi_t J^{L,\alpha}_t + (1-\pi_t) J^{L,SF}_t + \pi_{t+1} J^{L,\alpha}_{t+1} + (1-\pi_{t+1}) J^{L,SF}_{t+1} =
$$

$$
= \pi_t H(\alpha_t) - (1-\pi_t)\frac{U^{*2}}{4} + \pi_{t+1} H(\alpha_{t+1}) - (1-\pi_{t+1})\frac{U^{*2}}{4}
$$

s.t. reinforcement learning for π.
In the latter expression we made these substitutions:
$J^{L,\alpha} = H(\alpha) = \frac{-1}{2}(1-\alpha)^2 U^{*2}$ is payoff received by government in OC game; $J^{L,SF} = \frac{-U^{*2}}{4}$ is payoff received by government from PAL game.

Taking derivative with respect to α_{t+1}, we find that it is optimal to set α_{t+1} equal 1. Substituting that, we can find that α_t is the solution to:

$$
\alpha^* = arg \max_\alpha \pi_t H(\alpha_t) + (\pi_t + \pi_{t+1})\frac{U^{*2}}{4} \tag{34}
$$

The tradeoff is between obtaining good payoff in current period t by setting high α_t and obtaining high payoff in the next period $t+1$ taking that α_{t+1} will be set to 1. In order to be able to obtain high payoff at t+1, π_{t+1} must be high, and for this α_t must be low. So in each period t government solves for optimal α_t. From OC game with weighted by α payoff, optimal announcement is a function of α, optimal actual inflation is zero. Thus, government by setting α influences the announcement of inflation through value of which it tries to induce agent to play OC. Announcement should have appropriate value such that agent gets higher payoff from playing OC.

The results of simulations done by Deissenberg and Gonzales (2002) can be summarized as follows:

- α converges to 0.587: it is lower than $\alpha_2 = 0.88$ that makes agent indifferent between playing OC and PAL.

- π converges to 1: agent learns to play OC.

- Payoffs increase for both players and are better than under Nash or SF.

2 Description of the simulations.

In this section, we will present the description of the simulations that we performed. The structure of the simulations is represented in Figure 2. Briefly, the simulations have the following timing. First, government chooses selfishness parameter α and announcement y^a. Then agent chooses probability π, probability to play Optimal Cheating (OC) action, and the action of agent is determined randomly with this probability. Having observed agent's action, government chooses its actual inflation that depends on the strategy of the agent: there are two distinct levels of actual inflation for each agent's action. This is the end of the period, actual payoffs are computed for both agent and government. Then the sets of rules are updated with mutation and imitation, and the simulation continues to the next period. Two ways to compute payoffs were tried: simple payoff computed as in Deissenberg, Gonzales (2002) and weighted by the probability π (expected) payoff. This will be explained below in more details. Agent and government have their pools of rules. The initial sets of rules are generated randomly, and the decisions in the first period are random numbers. The decision rules in each period are chosen randomly based on rules' payoffs: the probability to be selected is higher for the rule with higher payoff. Now we will describe simulations with more details.

There are two types of decision-makers in our simulations: government and agent. Now we will describe their decision variables and their actions.

Government. The decision variables of government are α, selfishness parameter, and y^α, announcement of inflation. At the beginning of each period, government makes announcement of inflation y^α that may not be respected and decides how selfish α it wants to be.

Agent. The decision variable of agent is π, probability to play OC. Agent can choose whether to believe or not to believe the announcement. At the beginning of each period, agent chooses value of π. The decision "to believe" is made randomly with probability π, the decision "not to believe" is made, accordingly, with probability $1 - \pi$. The former is called optimal cheating action (OC); the latter is called private agent leadership (PAL): it is Stackelberg game with agent as a leader. To determine the decision, a random number is drawn from uniform distribution in interval [0,1]. If the value of the number drawn rn is less or equal to π, then OC is played, if otherwise, PAL is played. There are two distinct levels of inflation forecast that correspond to each action. If agent plays OC, the inflation forecast is set at:

$$x = y^a \tag{35}$$

which is determined from minimization of the agent's payoff function when agent believes that $y = y^a$.
If agent plays PAL, the inflation forecast is set at:

$$x = \frac{\theta(1-\theta^2)U^*}{1+\theta^4}, \tag{36}$$

which is the solution for forecast from the Stackelberg game with agent as a leader.

Government. It observes agent's decision of inflation forecast, and determines actual inflation based on this information. For each agent's decision, government makes different decision about actual inflation. If agent has played OC, then government sets actual inflation at

$$y = \frac{\alpha_{current}\theta U^* + x(1-\alpha_{current}+\alpha_{current}\theta^2)}{2-\alpha_{current}(1-\theta^2)} \tag{37}$$

which is derived from government's minimization of weighted loss function J^α taking α and agent's forecast as given.
If the agent has played PAL, government sets actual inflation at:

$$y = \frac{\theta U^*}{1+\theta^2} + \frac{x\theta^2}{1+\theta^2}, \tag{38}$$

which is the solution for actual inflation from the Stackelberg game with the agent as a leader.
This is the end of period, and both agent and government calculate their payoffs. Actual agent's payoff is computed as:

$$J^F = -0.5((y-x)^2 + y^2) \tag{39}$$

Government's payoff is calculated as weighted average of its own utility and agent's utility with α as a weight:

$$J^\alpha = \alpha J^L + (1-\alpha)J^F, \tag{40}$$

where J^L is government's own utility:

$$J^L = -0.5((U^* - \theta(y-x))^2 + y^2) \tag{41}$$

In these formulas, x is agent's inflation forecast, y is actual inflation, $\theta = 1$.

Update of pools of rules.
To start next period, agent and government update their pools of rules. Both agent and government have pools of 100 rules. Update consists of mutation and imitation.

Agent. Each rule π_j, $j = 1...100$, is mutated with mutation probability *mprob*.

Government. Each pair of rules α_j and y^a_j, $j = 1...100$, is chosen for mutation with mutation probability *mprob*, and then with probability 0.5, one of variables α or y^a is mutated.

When rule is chosen for mutation, it is changed as:

$$newrule = oldrule + random * deviation, \tag{42}$$

where random is random number drawn from standard normal distribution, deviation is the deviation for rule that is mutated. Deviations are different for inflation announcement, π and α.

After mutation, we evaluate agent's and government's rules in the environment of the previous period. Thus, when we calculate payoff for agent's rules, we take government's actions as given, equal to previous period values; and vice versa, when we calculate payoff for government's pair of rules, we take agent's action as given from the previous period.

Update with simple payoffs.

Computation of agent's payoffs. For each agent's rule of π_j, taking the random number rn drawn in the previous period to determine believe/not believe decision, we determine which strategy would have been played if a particular rule π_j had been chosen in the previous period. We call this potential action. If $rn \leq \pi_j$, then potential action is OC; if otherwise, then potential action is PAL. And so for each rule π_j, we can determine potential inflation forecast x_j^{pot} for each potential action. If potential action is OC, then potential forecast is equal to previous period government's announcement of inflation; if potential action is PAL, the potential forecast is equal to PAL forecast value from (36). Government's actual inflation y is taken as given from previous period. Thus, the payoff for each agent's rule is calculated as:

$$J_j^F = -0.5((y - x_{pot,j})^2 + y^2) \tag{43}$$

Computation of government's payoffs. Government's rule consists of two decision variables: announcement of inflation y^a and α. Government's payoff directly depends on its choice of α and actual inflation. In each period, actual inflation is chosen as a response to agent's forecast. If PAL is played, actual inflation depends on neither y^a nor α; if OC is played, then actual inflation is a function of α and y^a because agent's forecast is equal to y^a (see 35, 37). For each pair (y_j^a, α_j), potential actual inflation is determined taking as given the action of agent from the previous period. If PAL was played in the previous period, announcement did not affect government's payoff as actual inflation did not depend on announcement. Therefore, potential actual inflation is the same for each pair of rules (α_j, y^a), and equal to PAL actual inflation from the previous period. If in the previous period agent played OC, actual inflation depended on α and announcement y^a. So for each pair of rules (α_j, y_j^a), we calculate potential actual inflation that could have been set if this pair (α_j, y_j^a) had been chosen in the previous period. We do this using the formula for actual inflation (37) and substituting announcement y^a for agent's forecast:

$$y_j^{pot} = \frac{U^* \alpha_j \theta + y_j^a (1 - \alpha_j + \alpha_j \theta^2)}{2 - \alpha_j (1 - \theta^2)} \tag{44}$$

Payoff for each pair (α_j, y_j^a) is calculated with potential actual inflation y_j^{pot} and agent's inflation forecast x from previous period as:

$$J^{F,pot} = -0.5((y_{pot,j} - x)^2 + y_{pot}^2 \tag{45}$$

$$J^{L,pot} = -0.5((U^* - \theta(y_{pot,j} - x))^2 + y_{pot,j}^2) \tag{46}$$

$$J_j^\alpha = \alpha_j J^{L,pot} + (1 - \alpha_j) J^{F,pot} \tag{47}$$

We also perform simulations where the payoffs for both government and agent are averaged over 2 periods. Thus, the rules of α, the announcement of inflation and π are chosen in 2 periods from the same pool of rules. After mutation, the payoff for each of these 2 periods is calculated as described above. The payoff for each rule is then computed as average of these two payoffs. This was done to approximate the government's 2-period expected payoff used in Deissenberg, Gonzales (2002).

Update with weighted/expected payoffs.
The other way to compute payoffs is to use expected payoff, i.e. weighted average of OC and PAL potential payoffs with π as a weight. This way of payoff computation **insert explanation** So after the pool of rules has been mutated, payoff for each rule is calculated as follows.

Computation of agent's payoffs. For agent, there are 2 distinct potential inflation forecasts. Potential OC forecast is equal to the announcement of inflation from the previous period; potential PAL forecast is calculated as forecast from private agent leadership (36). Potential OC payoff is computed using potential OC forecast, potential PAL payoff is computed using potential PAL forecast. The actual inflation is taken from previous period. Then for each rule of π_j, these two payoffs are weighted by π_j:

$$\begin{aligned} J_j^F &= \pi_j J_{pot}^{F,OCA} + (1 - \pi_j) J_{pot}^{F,PAL} = & (48) \\ &= -0.5(\pi_j((x_{pot,oca} - y)^2 + y^2) + (1 - \pi_j)((x_{pot,pal} - y)^2 + y^2) & (49) \end{aligned}$$

Computation of government's payoffs. For government, there are 2 distinct potential actual inflation levels: OC potential value calculated for each pair of rules from (44) and PAL value from (38). Then payoff for each pair of rules is calculated as weighted average of OC and PAL potential payoffs where weight is value of π from previous period:

$$J_j^\alpha = \pi J_{pot}^{\alpha,OCA} + (1 - \pi) J_{pot}^{\alpha,PAL}, \tag{50}$$

where $J_{pot}^{\alpha,OCA}$ is potential OC payoff calculated as (47) with y^{pot} from (44), and $J_{pot}^{\alpha,PAL}$ is potential PAL payoff calculated as (47) with y^{pot} from (38).

After mutation, imitation is done by tournament selection. For each rule or pair of rules j, the following procedure is done: 2 rules are drawn randomly from the pool of rules, and rule j is replaced by the one with higher utility. We also performed simulations with imitation by roulette wheel. Two ways of imitation did not bring any difference in results. Having updated their pools, agent and government make new decisions in the next period. This is done randomly with the probability for the rule to be selected based on the rule's payoff: the higher the payoff, the higher the probability for the rule to be chosen. The probability is computed as:

$$P(rule_j) = \frac{payoff_j}{\sum_{i=1}^{N} payoffi} \tag{51}$$

We have done simulations with different values of parameters. The different values of the parameters are given in the following table.

parameter	value 1	value 2
mutation probability	0.033	0.330
deviation of announcement	0.1	3.0
order of mutation/imitation	mutation first	imitation first
payoff computation	simple	weighted

3 Results of simulations.

We will describe our results for two groups of simulations - with simple and weighted payoffs.

3.1 Simulations with simple payoffs.

The following regularities were observed in the simulations with simple payoffs. Selfishness parameter α converges to 0 or very small values close to 0. Probability to play OC π fluctuates between 0 and 1, staying at 0 or 1 for variable lengths of time that depend on the simulation type. Announcement is volatile, and when it converges, the value to which it converges vary for different simulations. When π goes to 1, announcement is always below 5. We present time series from simulations in Figures 3,4 and Figures 6,7. The frequency data is shown in Figures 5 and 8. Next we provide our explanation for these observations.

The selfishness parameter α takes values around zero because government is always indifferent or better off being absolutely altruistic. If PAL is played, we can compute agent's and government's payoff by substituting agent's forecast (36) and government's actual inflation (38) into their payoff functions. And we find that their payoffs are equal to -6.25. If PAL is played, value of α does not matter for the computation of the payoffs. If OC is played, and this happens when announcement is below 5, government's payoff J^L is always lower than agent's payoff J^F (both numbers are negative). Therefore, J^α is higher when government sets α equal zero. Thus, government becomes absolutely altruistic in order to maximize its payoff.

π converges to either zero or one and fluctuates from one of these values to the other depending on the values of α and announcement y^a. If announcement is higher than 5, then agents are better off playing PAL (not believe); and this brings forecast equal 0 and actual inflation equal 2.5. If announcement is below 5, then agent's π is high approaching 1. Then forecast is equal to announcement from equation (35), and actual inflation is approximately equal to half the announcement as α is about 0 from equation (37).

When announcement converges to some level, this level is indeterminate and can be anything. If the value to which announcement converges is below 5, π goes to 1. If announcement sets at some value above 5, π approaches 0. To determine the threshold level for the switch of π, we need to find the condition when agent's payoff from playing PAL is less or equal to payoff from playing OC. We do this by substituting values of forecast (36),(35) and actual inflation (38),(37) into agent's payoff function with parameter $\theta = 1$:

$$J^{F,PAL} \leq J^{F,OCA} \tag{52}$$

$$-\frac{1}{2}[(0-U^*)^2+(U^*)^2] \leq -\frac{1}{2}[(y^a-(\alpha U^*+\frac{y^a}{2}))^2+(\alpha U^*+\frac{y^a}{2})^2] \tag{53}$$

Rearranging, we can get condition for agent to play OC:

$$(\frac{U^*}{2})^2 \geq (\alpha U^*)^2+(\frac{y^a}{2})^2 \tag{54}$$

For $\alpha = 0$ (as α converges to 0 in simulations), this condition becomes:

$$-U^* \leq y^a \leq U^* \tag{55}$$

So if y^a is in this range, payoff from playing OC is higher and, therefore, agent wants to play it and sets π close to 1. With constant probability of mutation and especially with high deviation of inflation announcement, several changes of announcement are possible during each simulation. Once the announcement crosses the threshold value of 5, this brings subsequent change in π.

For high value of deviation and mutation rate, the resulting time series are more volatile. The volatility is driven by the changes in announcement and α. As government's announcement changes to high values above 5, agents are worse off if they play OC, and their π drops to low around zero. Or if government's α changes to value higher than 0 (around 0.2-0.3), agents change their π as agent gets lower payoff with higher value of α. As a result of these changes in announcement and $alpha$, we can observe π fluctuating between 0 and 1. This volatility is caused by changes in announcement more frequently than by changes in α.

Simulations with low deviation of announcement have lower volatility. The results seem to be drawn by the announcement value. So if at the beginning of the simulation, announcement sets at higher than 5, π goes to lower values, close to 0. If deviation is low, the opportunity for announcement to change by relatively big amount and cross the line of 5, and so influence π to go from 1 to 0 or the other way around, is low. Therefore, these simulations have less volatility than those with high deviation.

Higher mutation rate brings more frequent volatility to time series. High deviation brings higher amplitude of changes: swings in announcement are bigger. For lower mutation rates and lower deviation, it takes more time for changes to happen because when changes happen they are very small at one time.

Simulations with imitation first show the same regularities as simulations with mutation first but with more volatility. This happens because, in this case, imitation does not eliminate rules resulting from mutation that are very different from the rest of the rules in the pool and may, therefore, have lower payoff, but they get chosen as the choice is random though based on payoffs.

Simulations with simple payoffs averaged over 2 periods showed the same results as simulations with simple payoffs. The times series are depicted in Figures 12,13, and frequency data is on Figure 14.

More detailed analysis of observations. We analyzed changes in π, announcement and α. In time series, we can observe that π can change from 1 to 0, and from 0 to 1. Usually, these changes happen simultaneously with changes in α and announcement. It is interesting to analyze which of these variables trigger the change, and how it happens. It seems that changes in π are usually triggered by changes in announcement or α or both.

We analyzed change of π from 1 to 0 in the simulation with high deviation and high mutation (it is depicted in Figure 3; sim2). Up until and including period 794, α settles at 0, and announcement settles at 3.8794. π is 1 or very close to 1. So, OC is played with rare accidents of PAL (because π is not exactly equal to 1, and choice is random). And so in 794, PAL is played. In period 795, government evaluates its rules in the environment of period 794. As in period 794 neither α nor announcement were used to compute actual inflation, all government's rules have the same potential value of actual inflation equal to PAL actual inflation (38) and get the same potential payoff and, therefore, equal probability to be chosen in the next period. As a result, in period 795 announcement is -4.8322, α is still 0. This becomes the turning point. Now government wants to choose negative announcements as they give higher payoffs than positive ones. In 796 announcement is chosen to be -5 although announcement -4.8322 gives higher potential payoff, however, the choice is random. In 797, government chooses positive α of 0.1893, announcement of -4.8322; this rule gives the highest payoff among all available rules. When announcement is negative, rules with positive α get higher payoff than rules with zero α. As α increases to positive value, agent becomes better off playing PAL. This happens because the deviation of potential PAL forecast of 0 from actual inflation is smaller than that of potential OC forecast. This happens because OC actual inflation includes term αU^* (see 37), and when announcement is negative, adding positive term increases actual inflation and reduces deviation term from potential PAL inflation of 0. And so low values of π give higher potential payoffs and start to dominate in the pool of rules.

This completes the explanation of transition of π from 1 to 0. So even though announcement is within range where PAL payoff is lower than OC payoff, high value of α makes agents better off by playing PAL.

It is interesting to address the question how the reverse change happens: how π changes from 0 to 1. This is done for simulation in Figure 3. Up until period 394, π settles at 0 or close to 0 values, announcement - at around 15, α - at 0.0147. So agent plays PAL, and when PAL is played, the computation of potential payoffs does not include announcement and α, and so all governments rules receive the same payoff and equal probability to be chosen. Thus, accidentally, rule with lower announcement gets selected. At this point, agents π is very low, close to zero, but not zero. So accidentally low random number is drawn, and OC is played. When this coincides with governments experiment with low announcement, both government and agent start to learn about new mode of behavior: low announcement and high π, respectively. Low value of announcement is essential to convince agent to believe. If α is zero, but announcement is higher than 5, agent receives higher payoff from PAL than from OC. So it is very important that announcement decreases below 5 for π to go up to 1.

We can observe that in time series, the announcement changes to low values first, then α decreases to low zero value. Agents payoff is always negatively related to α, so when α increases, agents OC payoff goes down:

$$\frac{\partial J^{F,OCA}}{\partial \alpha} = -2\alpha U^*, \tag{56}$$

which is always negative. So when α goes down, agents OC payoff increases. This explains cases when α increases, π goes down, and the other way around. But the first change to induce switch of π from 0 to 1 is change in announcement.

To summarize, the change in π from 1 to 0 can be triggered by either increase of announcement above threshold level of 5 or by increase in α above 0; and change in π from 0 to 1 is always triggered by decrease of announcement below 5.

Explanation of why announcement does not go to 0. When α goes to 0, government's payoff is equal to agent's payoff, and its maximum value is achieved with announcement equal to 0. Because when OC is played, agent's forecast is equal to announcement, and then actual inflation is equal to half announcement, and then both agent's and government's payoffs are equal to 0. The values of forecast, actual inflation and unemployment are equal to those in Stackelberg with government as a leader (Ramsey outcome).

However, in our simulations we did not observe convergence of announcement to 0. When announcement is lower than 5, it can take any value, and as long as it is below 5, and α is zero, π is close to 1. We will analyze this question using simulation from Figure 3. In period 401, announcement starts to decrease, and by period 430, it settles at value 1.5056 for 4 consecutive periods. In period 431, announcement 2.1399 is played, in period 434, announcement 3.2303 is

chosen accidentally because these rules have lower payoffs in the corresponding pools of rules from which they were chosen. And although rule with announcement 1.5056 has lower payoff than rule with announcement 2.1399 (or 3.2303), in the next period when it is evaluated in the environment of previous period when announcement was 2.1399 (3.2303). However, pools of rules are already dominated by rules with α zero and announcement 1.5056, so that it gets chosen and further replaces other rules. Starting from period 435, announcement and α settled at values 1.5056 and 0, and continue to stay at these values till the end of 1000-period simulation. Occasionally, PAL is played, or α is chosen to be non-zero, but very close to zero. When we look at rules and their payoffs, we can see that, for example, in period 447 rule with announcement 1.0024 has lower payoff than rule with announcement 1.5056, with α being 0 for both of them. This can be explained by how payoff is calculated. Agent's forecast is taken from previous period, and it is equal to previous period announcement that was 1.5056. For announcement 1.0024, potential actual inflation (from 44) 0.5012. Then this rule's payoff is calculated by (47) and is equal to $\frac{-1}{2}((1.5056 - 0.5012)^2 + 0.5012^2) = -0.6300$
The potential payoff of rule with announcement 1.5056 is equal to:
$\frac{-1}{/}2((1.5056 - 0.7528)^2 + 0.7528^2) = -0.5667$

Thus rule with higher announcement gets higher payoff because that is the value of announcement that was chosen in the previous period. Similarly, in period 454 rule with announcement 0.7946 gets lower payoff than rule with announcement 1.5056. As a result, when announcement settles at some value, which is initially randomly determined, then it is difficult for government to choose any announcement that differs from this value substantially as this increases deviation from the forecast squared term. Therefore, the best value of zero announcement might never be achieved or needs to happen very slowly. So even though lower announcement can give higher absolute payoff, the rules are evaluated in the environment of the previous period, and the relative value of payoff of each rule in the pool of rules matters. And the lower this payoff is, the higher is the difference between previous period announcement and the announcement of each rule.

3.2 Simulations with weighted payoffs.

These simulations show the following regularities. α converges to 0 or values very close to 0. Announcement converges to 0 from any initial value and stays there for the rest of simulation with some volatility coming from mutation. Probability to play OC π goes to 1 or values very close to 1, and stays there most of the time with occasional and short drops to 0. The periods when π is around 0 are much shorter than those in the simulations with simple payoffs. It was also observed that drops in π are initiated by changes in π itself, unlike in the simple simulations where π changed as a reaction to changes in α or announcement or both. See Figures 9,10 for graphs with time series from simulations with weighted payoffs., frequency data is presented in Figure 11.

Time series and pools of rules were analyzed during periods of change of π from 1 to 0, and from 0 to 1. This was done for the simulation from Figure 5. In this simulation, π changed first, and then changes in announcement and α followed. The course of events was the following. Announcement and α stay for some time at values -0.0043 and 0 respectively; π was very close to 1 or equal 1. In period 243, the random number turns out to be lower than current value of π, and PAL is played. In the next period, rules are evaluated taking forecast and actual inflation equal 0 and 2.5 respectively. Payoff for each rule is weighed average of OC and PAL potential payoffs. OC potential inflation is equal -0.0043 that is lower than PAL value of 0. Potential OC payoff is lower than PAL payoff because discrepancy between potential inflation of -0.0043 and actual inflation 2.5 from the previous period is higher than discrepancy between 0 and 2.5 for PAL. The weight used for weighting payoff is π, and so rules with lower π have higher payoff because they give less weight to more negative potential OC value. This is the way through which the rules of lower π survive in the population and eventually dominate. It takes 10 periods for agents to switch to very low and then zero values of π.

It is interesting to observe that government does not change its decision until period 266, when α is chosen to equal 0.1745. This happens because in the previous period, 265, PAL is played, and π is equal to 0, and this value serves as a weight in payoffs. So when government's rules are evaluated in the next period, zero weight is given to OC potential payoff. In addition, all of the rules receive the same PAL payoff of -6.25, because values of announcement and α are not accounted for in PAL computations. And so high value of $\alpha = 0.1745$ is chosen because all rules have equal payoffs. Subsequently in period 268, when π is 0.012, rules with higher α have lower payoffs, and a rule with lower $\alpha = 0.0906$ was chosen. So when agent chooses π equal 0, in the next period each government's rule of announcement and α has equal probability to be chosen. This can bring the next change of π from 0 to 1. It is interesting to observe, for example, in period 276, the value of announcement is chosen to be 2.3826, which is very close to PAL actual inflation of 2.5. When rules of π are evaluated in the next period, potential OC payoff is higher than potential PAL payoff because discrepancy between actual inflation 2.5 and potential forecast 2.3826 is smaller than between PAL values of 2.5 and 0. Thus, higher rules of π get higher payoff because they give more weight to higher OC potential payoff. Thus, the evaluation of π rules highly depends on the circumstances of the previous period. However, even though agent starts to experiment with higher π, 2 periods of non-zero π, 277 and 278, are not enough to have impact on government's behavior. However, during the next 20 periods government learns to use lower values of α and announcement. It is interesting to see whether this change in government's behavior brings change to agent's choice of π. To summarize, this transition from 1 to 0 happens when, by chance, PAL is played, and OC forecast is lower than PAL zero forecast which gives higher payoff to lower π rules.

The next question to be analyzed is transition of π from 0 to 1. We do this in the context of the same simulation.

When agent chooses π equal 0, in the next period when government's rules are evaluated, they all get the same payoff equal to PAL value of -6.25, as weight on potential OC payoff is equal to zero. This way, government gives equal probability to all rules when next decision is made. Only when π is non-zero, rules of announcement and α are evaluated.

Government's payoff depends on potential OC and PAL values. Contribution from potential PAL payoff is the same for all government's rules, as potential actual inflation is the same and equal to 2.5. Thus, differentiation in payoffs for government's rules comes from OC potential payoff. Those rules that have lower potential actual inflation and/or value of potential actual inflation closer to agent's forecast from the previous period have higher payoff. It was already argued above that it is always better decision for government to be absolutely altruistic, and have low, close to zero, α. It can be observed that government learnt to set low values of α, close to 0, during 10 periods (since 488 to 498).

It is more difficult to learn about announcement. For potential payoff to be low, potential actual inflation must be low to minimize term y^2, and also potential actual inflation must be close to forecast from previous period to minimize the term $(x-y)^2$. As forecast is equal to 0 in PAL, and is equal to announcement for OC, and there is randomness in this game, it takes longer to adjust and set announcement correctly. It must also be noted that if PAL is played and π is non-zero in period t, this gives government very strong incentives to set announcement to 0 in period $t+1$. If this happens, payoff is maximized at 0. If OC is played at t, then there is trade-off between minimizing y^2 and discrepancy squared terms, therefore [I think], we do not observe perfect convergence of announcement to 0.

When chosen π is non-zero, this triggers the above described changes in government's behavior that, in its turn, affects agent's π. When government has low announcement, and if PAL is played in the previous period, higher rules of π have higher payoffs because they give more weight to higher potential OC payoff. It is higher because discrepancy term is lower, as potential forecast is equal to announcement, and it is non-zero positive value starting from period 494 with occasionally low negative values. In this way, rules with higher π get selected. When OC was played in the previous period, then actual inflation is low. For example, starting from period 527, actual inflation is 0.099. Announcement is equal 0.1981. When rules of π are assessed, potential PAL payoff is $\frac{-1}{2}[(0-0.099)^2+0.099^2]=-0.0010$. Potential OC payoff is $\frac{-1}{2}[(0.1981-0.099)^2+0.099^2)=-0.0010$. They are equal because $\alpha=0$, and so actual inflation is equal to half of announcement. This gives us 2 results. First, it is the explanation of selection of higher π. When PAL is played, rules of higher π get over the population. When OC is played, this trend in changing π population is not reversed, although not strengthened either. So rules with higher π dominate in the pool of rules. Second, this gives explanation of non-convergence of announcement to 0. Because, α is 0, government's payoff is equal to agent's payoff. And potential OC and PAL values are the same, there is no incentive to change announcement.

It was observed that the magnitude of the payoffs of the rules for both agent and government depend largely on the value of announcement and actual inflation from the previous period. Explain relative performance of rules. Talk about better interplay between variables in weighed model.

3.3 Discussion

.

As it was explained above, Deissenberg and Gonzales do the following in their paper. In each period government solves for optimal α_t, taking α_{t+1} equal to 1. From OC game with weighted by α payoff, optimal announcement is a function of α, optimal actual inflation is zero. Thus, government by setting α influences announcement through value of which it tries to induce agent to play OC. Announcement should be proper value such that agent gets higher payoff from playing OC. In their simulations, they observed convergence of π to 1, α to 0.587.

In our simulations, we obtained fluctuations of π between 0 and 1 depending on the value of announcement and α. In our case, α and announcement are not directly connected, they are chosen from the respective ranges randomly and undergo mutation. This brings recurrent changes in announcement and α, and, as a result, π changes too. Therefore, we do not observe final convergence. There is always possibility for opposite change.

We can also interpret the obtained results from the simulations as an explanation for the change in average inflation over time. Once government starts to experiment with the announcement of inflation or selfishness parameter, and sets announcement higher than threshold level of 5 or increases α above 0, it becomes optimal for agent to respond by playing PAL and not to continue to play OC. The increase of government's announcement or selfishness parameter leads to agent's play of PAL, and consequently, actual inflation increases to PAL value of 2.5. The decrease in average inflation can be explained by the following behavior of the government. When government reduces its announcement of inflation below threshold level of and is altruistic ($\alpha = 0$), it becomes optimal for agent to respond with playing OC. When announcement is below 5, $\alpha = 0$, the actual inflation is equal to half of announcement. And so the reduction in government's announcement of inflation triggers play of OC and decline in actual inflation.

4 Conclusion.

In this model, there is a strong link between the announcement and outcome of actual inflation. This means that private agent wants to take announcement into consideration as it communicates information about future government's action. We found that agent's willingness to believe government depends on the value of the announcement. If it is lower than threshold level, then agent believes announcement and is cheated upon; if it is higher, then agent acts as if s/he does not believe. Government's action depends on agent's decision. Government's best solution is to announce zero inflation in this setup, and if it does, agent is better off to believe. However, this does not always happen. And the reason is not the same as in the models with asymmetric information where zero-inflation policy is not credible. In our simulations, government randomly coordinates on some low level of inflation and cannot learn zero-announcement because it is short-sighted due to the setup of update procedure. The results of our simulations are different from those by Deissenberg and Gonzales because of the differences in the learning abilities that government and agent are endowed with. In Deissenberg and Gonzales (2002), the probability to believe converges to 1. In our simulations, the probability to believe fluctuates between 0 and 1 depending on the values of announcement and selfishness parameter. And as these variables change due to mutation, the optimal decision by agent to believe/not believe can change to the opposite. This also brings changes in actual inflation, and so this model can be used to explain changes in average inflation over time.

References.

Arifovic, J., Ledyard, J. 2004. Computer Testbeds: The Dynamics of Groves-Ledyard Mechanisms. Manuscript, April.

Backus D.K., Driffill J. 1985. Inflation and Reputation. *American Economic Review*, 75, no.3 (June): 530-538.

Ball L. 1995. Time Consistent Inflation Policy and Persistent Changes in Inflation. *Journal of Monetary Economics*, 36, no.2 (November): 329-350.

Barro R.J., Gordon D.B. 1983. Rules, Discretion, and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics*, 12, no.1: 101-121.

Cukierman A., Liviatan N. 1991. Optimal Accommodation by Strong Policymakers under Incomplete Information. *Journal of Monetary Economics*, 27, no.1 (February): 99-127.

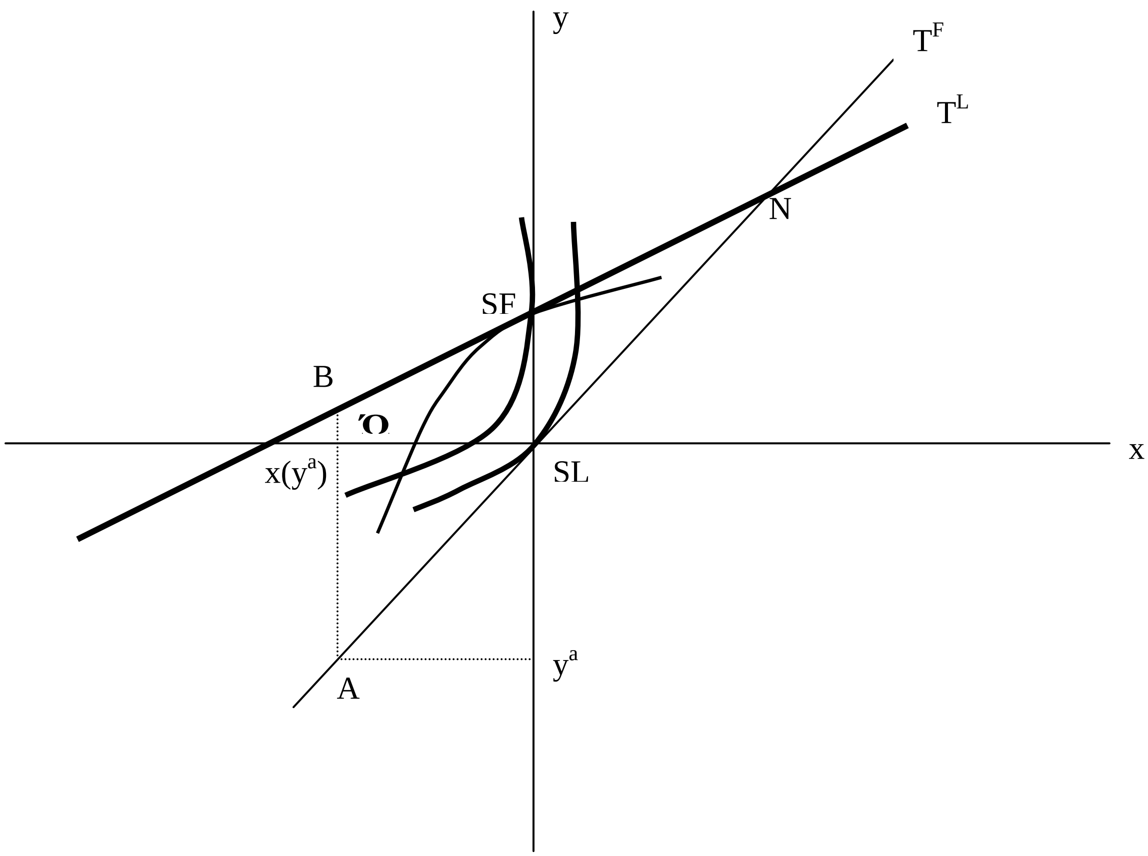
Deissenberg C., Gonzales F.A. 2002. Cheating for the common good in a macroeconomic policy game. *Journal of Economic Dynamics and Control*, 26, Issues 9-10 (August): 1353-1785.

Kydland F.E., Prescott E.C. 1977. Rules Rather Than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, 85,no.3 (June): 473-491.

Sargent Th.J. 1999. *The Conquest of American Inflation.* Princeton, New Jersey: Princeton University Press.

Stein J. 1989. Cheap Talk and the Fed: A Theory of Imprecise Policy Announcement. *The American Economic Review*, 79, no.1 (March): 32-42.

Figure 1. The solutions of the Kydland-Prescott model.



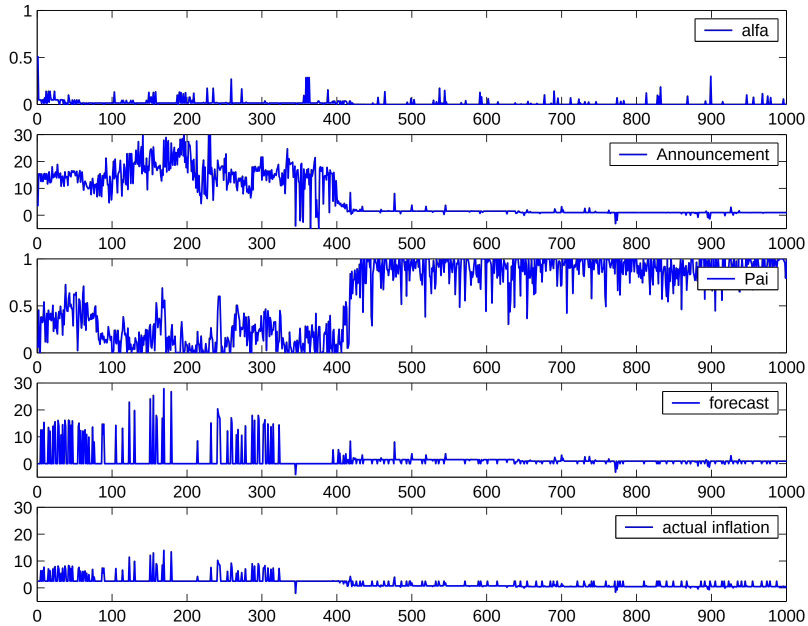


Figure 3. Simulation 1 with high deviation, high mutation and simple payoff.

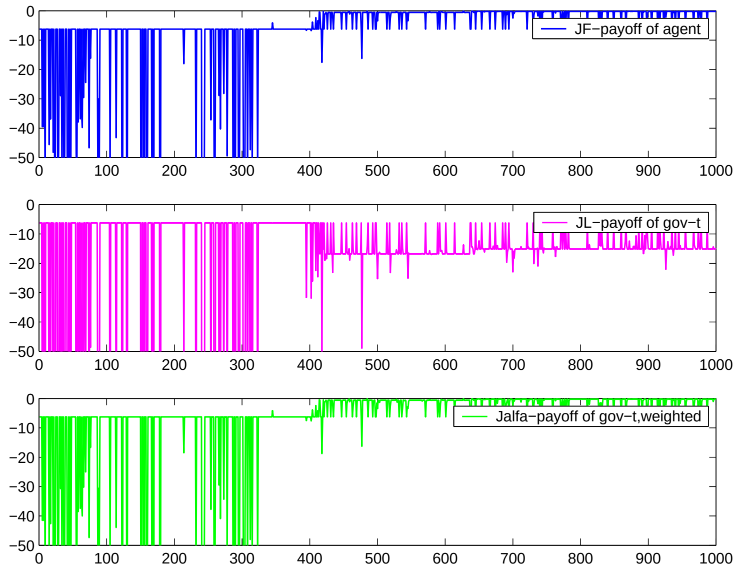


Figure 4. Payoffs for simulation 1 with high deviation, high mutation and simple payoff.

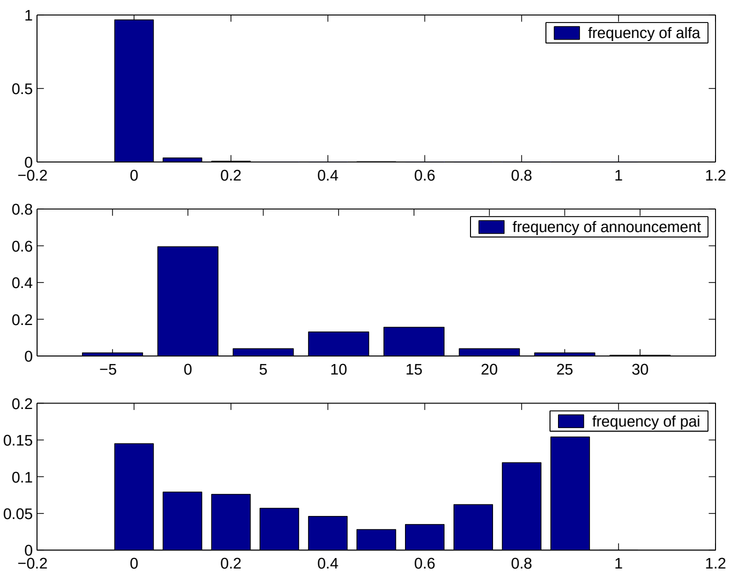


Figure 5. Frequencies for simulation 1 with high deviation, high mutation and simple payoff.

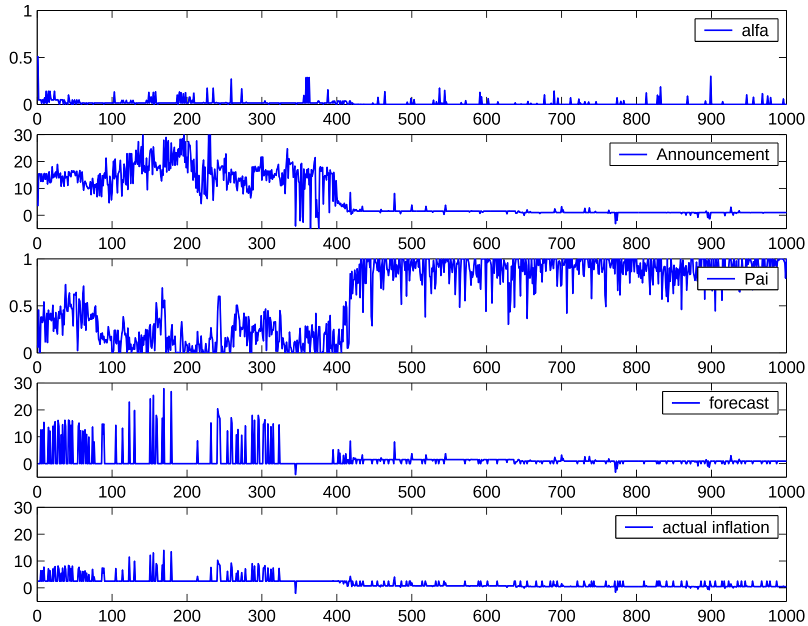


Figure 6. Simulation 2 with high deviation, high mutation and simple payoff.

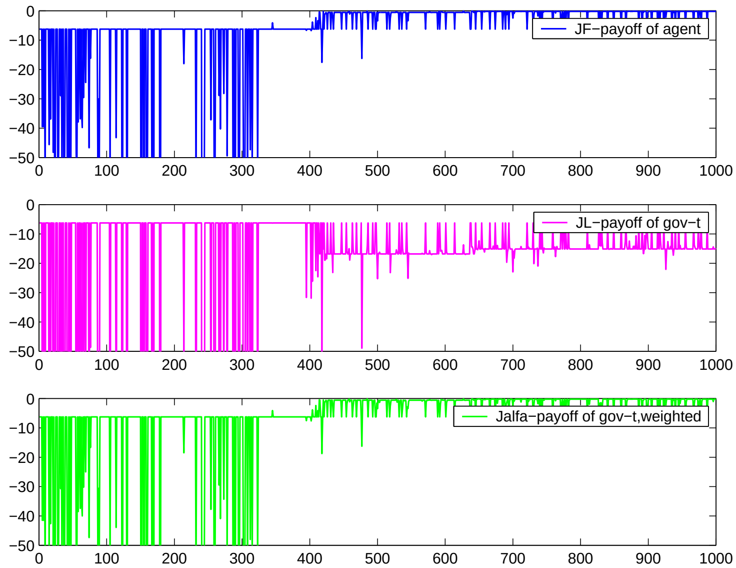


Figure 7. Payoffs for simulation 2 with high deviation, high mutation and simple payoff.

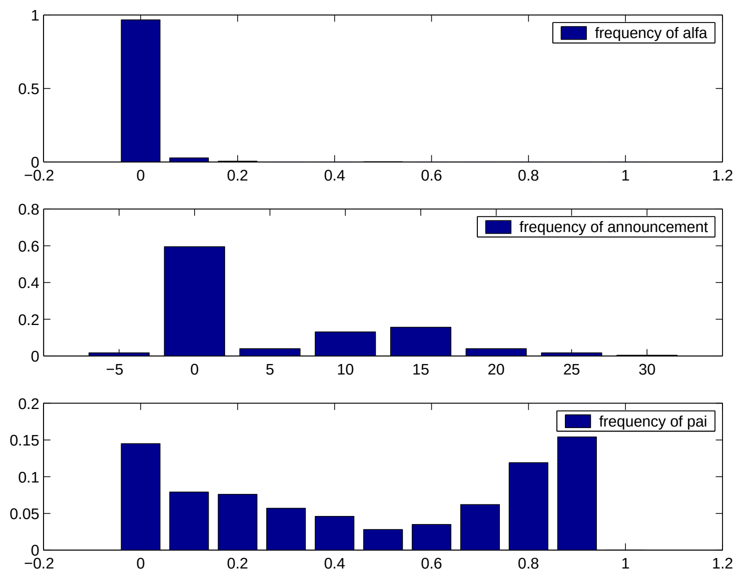


Figure 8. Frequencies for simulation 2 with high deviation, high mutation and simple payoff.

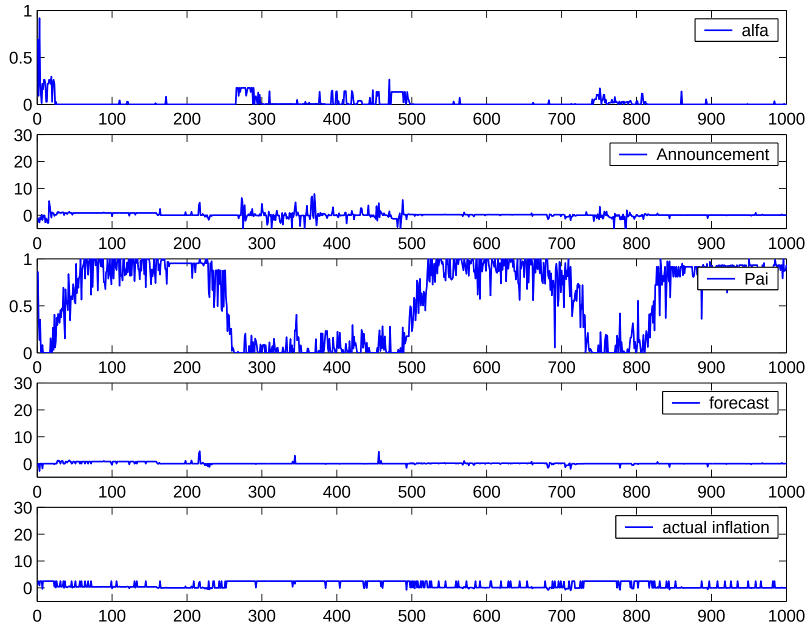


Figure 9. Simulation with high deviation, high mutation and weighted payoff.

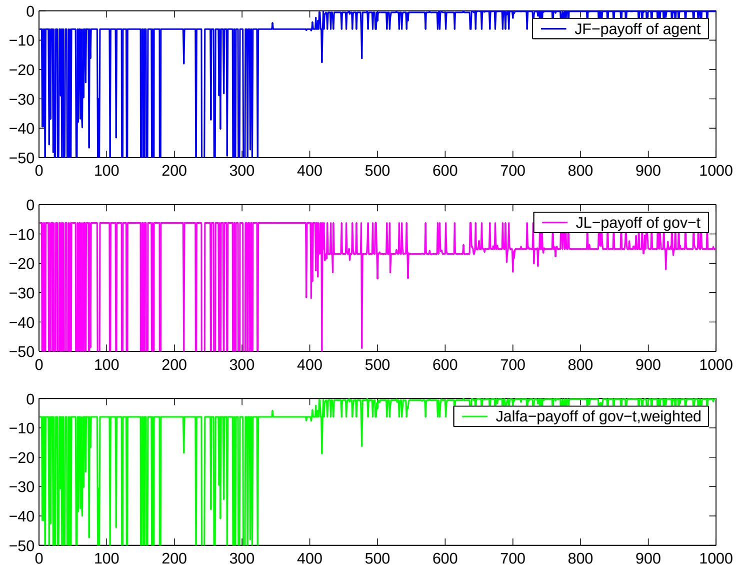


Figure 10. Payoffs for simulation with high deviation, high mutation and simple payoff.

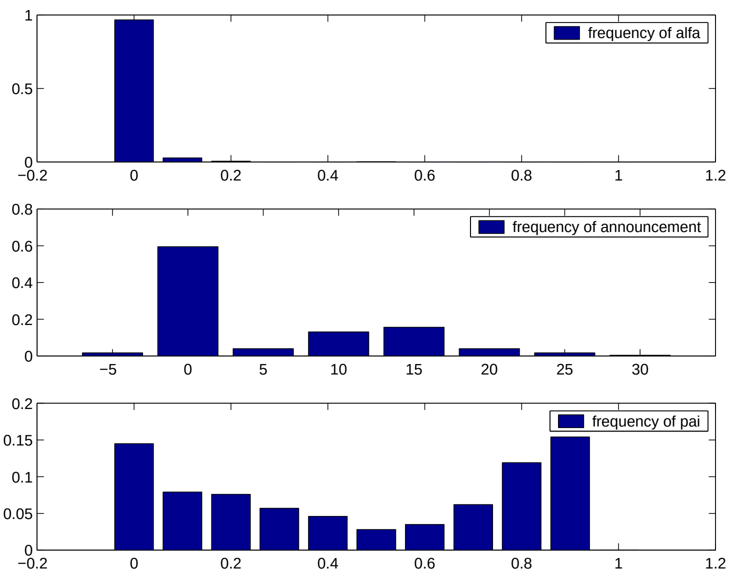


Figure 11. Frequencies for simulation with high deviation, high mutation and simple payoff.

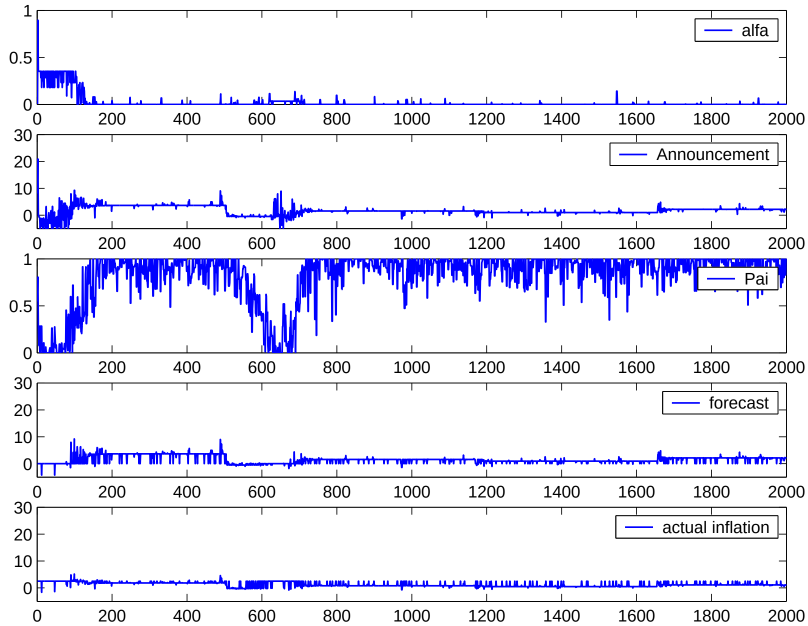
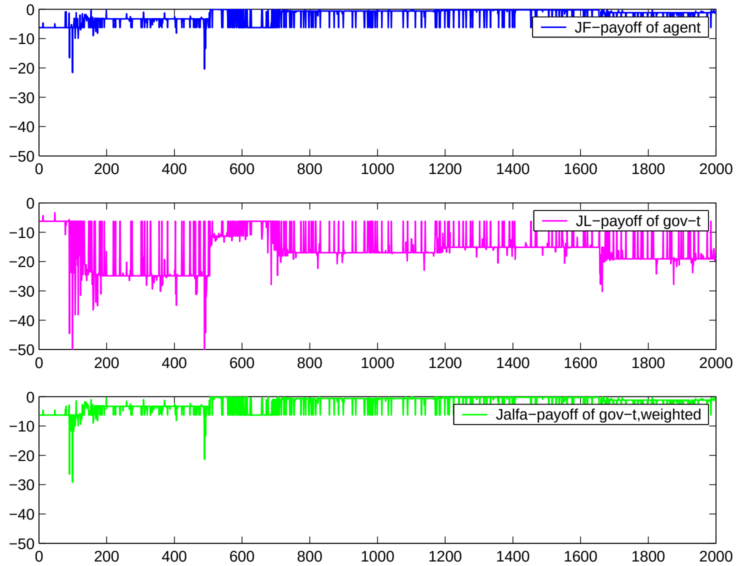
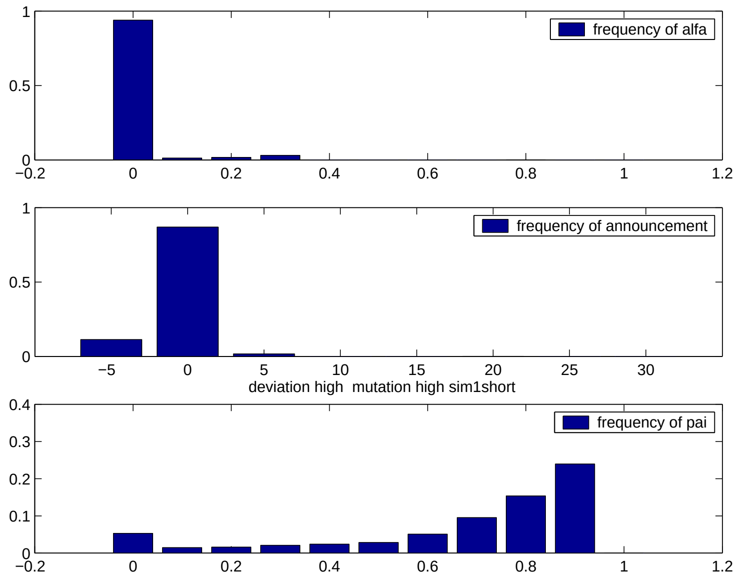


Figure 12. Simulation with high deviation, high mutation and simple payoff averaged over 2 periods.



ure 13. Payoffs for simulation with high deviation, high mutation and simple payoff averaged over 2 per



e 14. Frequencies for simulation with high deviation, high mutation and simple payoff averaged over 2 p