*Games* **2014**, *5*, 1-25; doi:10.3390/g5010001

OPEN ACCESS

*games*

**ISSN 2073-4336**

www.mdpi.com/journal/games

*Article*

# Introducing Disappointment Dynamics and Comparing Behaviors in Evolutionary Games: Some Simulation Results

**Tassos Patokos**

University of Hertfordshire, Hertfordshire Business School, Department of Accounting, Finance and Economics, Hatfield, AL10 9AB, UK; E-Mail: t.patokos@herts.ac.uk

*Received: 26 November 2014; in revised form: 11 January 2014 / Accepted: 23 January 2014 / Published: 30 January 2014*

---

**Abstract:** The paper presents an evolutionary model, based on the assumption that agents may revise their current strategies if they previously failed to attain the maximum level of potential payoffs. We offer three versions of this reflexive mechanism, each one of which describes a distinct type: spontaneous agents, rigid players, and 'satisficers'. We use simulations to examine the performance of these types. Agents who change their strategies relatively easily tend to perform better in coordination games, but antagonistic games generally lead to more favorable outcomes if the individuals only change their strategies when disappointment from previous rounds surpasses some predefined threshold.

**Keywords:** game theory; reinforcement learning; adaptive procedure; revision protocol; disappointment; simulations

---

## 1. Introduction

Individuals are averse to unpleasant experiences. When such experiences happen, it makes sense to assert that the individuals affected shall choose different strategies from those that brought about unsatisfactory outcomes. In this paper we present three variations of a learning model, which reflects the intuitive fact that agents try to eschew perceived disappointing outcomes. Disappointment emerges when a player fails to achieve the maximum level of payoffs that would be possible, had a different strategy been chosen. In other words, the core assumption is that a differential between someone's actual payoffs and the maximum level of potential payoffs (with the opponent's choice taken as a given) generates a tendency to choose a different strategy in the next round of the same game.

While it seems safe to conjecture that individuals avoid disappointment in general, individual reactions to past disappointment outcomes are contingent on psychological issues, and, as such, they may vary dramatically across persons. For example, a person with relatively low tolerance to disappointment might be expected to change their strategy after a disappointing outcome with higher probability than another person who is more patient. Evidently, the individual psychological profile is important in determining action. Each one of the three variations of the learning model we describe represents a distinct behavioral type: we deal with spontaneous and impatient agents, players who are rigid and display high inertia, and 'satisficers'. We use a simulation program to study the evolutionary equilibria in an assortment of 2 × 2 games with populations consisting of the aforementioned behavioral types, the aim being to compare these types' performances in coordination and antagonistic games.

Several prominent authors (such as Sugden [1] or Rubinstein [2]) have expressed the opinion that economics of bounded rationality does not need more theoretical models, but rather, must focus on empirical research. While we embrace this view, the underlying behavioral hypotheses we present here do not seek to explain specific empirical or experimental findings; however, our purpose is not to merely enrich the literature with more learning or adaptation rules, which would perhaps seem superfluous. Rather, we focus on providing a better handle on how a combination of limited computational power and of a psychological aversion to disappointment matters. Although the rules we use are *ad hoc* constructed, they serve as proxies of players' real-life attitudes (for example, patience or spontaneity). In fact, our simulation results show that, depending on the game, different behavioral types are bound to be more successful than others; this means that a single reflexive or adaptation model is most likely to be insufficient for studying interactions, which reinforces the need for more empirical research.

One might wonder why we choose to introduce newer learning rules, rather than use something from the wealth of rules that can be found in the literature. The three rules presented in this paper have a clear behavioral background, based on the postulation that what determines human action now is possible disappointment experienced in past play. Therefore, even if the adaptive procedures we suggest could be approximated by an existing rule, we are interested in exploring how the *specific* stylized assumptions (each one of which is linked to a psychological type) translate to a learning model, and at a second level, in comparing these particular rules as to their social efficiency; it is then possible to conclude if these characteristics help the individuals attain better outcomes (such as a peaceful resolution in the "Hawk–Dove" game, or mutual cooperation in the "Prisoners' Dilemma"). The corresponding findings, even if they are on a theoretical level, are important to those interested in the psychology of strategic behavior, in that they provide newer insights on intertemporal strategic play; and if the behavioral type can be seen as a control variable (i.e., if the individual may choose their own type), then this theoretical approach is apt to suggest what type would be preferable, given the nature of the interaction.

The paper is structured in five sections. Section 2 provides a concise review of the use of adaptive procedures in evolutionary game theory, along with the notion of disappointment in economics. Section 3 presents the three different agent type protocols. Section 4 discusses simulations with and without random perturbations (i.e., noise on the Markovian process), and Section 5 concludes.

## 2. Evolutionary Dynamics and the Notion of Disappointment

Evolutionary game theory took off in the first half of the 1970s, after Maynard Smith and Price [3] and Maynard Smith [4], inspired by Lewontin [5], applied game theory to biology and defined the evolutionary stable strategy. Taylor and Jonker [6] proposed replicator dynamics as a way to translate this process to mathematical language. Although very popular (mainly due to their simplicity), replicator dynamics are not generally thought of as particularly apt for economic applications[1] because they are seen as a highly restrictive selection model. Interest, therefore, shifted to dynamics that seek to describe how a population increases along with the success of its adopted strategy. According to Friedman [11], the most abstract way of modeling a selection process is to assume growth rates that are positively correlated with relative fitness.

Usually, evolutionary selection processes do not just portray how a population grows over time, but they are also based on an underlying series of assumptions that offer an explicit rationale of how agents adapt with historical time. Except for natural selection, Young (1998) distinguishes between rules of imitation (for example, [12]), reinforcement learning [13–15], and fictitious play (for example, [16]). In some of these models, beliefs on what the opponent plays are updated by means of Bayes' rule (see [17]).

Evolutionary dynamics can be deterministic or stochastic. Deterministic evolutionary dynamics (such as the replicator dynamics) describe the evolutionary path with systems of differential (or difference) equations; each one of these equations expresses the increasing (or decreasing) rate of some strategy's frequency as a function of the portion of the population who chooses this strategy, and in accordance to the assumed revision protocol [18–19]. In such models, mutations are thought of as rare and random, and therefore, not continual or correlated in any way. Contrary to this assumption, stochastic evolutionary dynamics examine the possibility that mutations cause enough "noise" to pull the population state out of a certain basin of attraction. These models incorporate perturbations by use of Markov processes and may lead to fundamentally different results than deterministic dynamics. Among the seminal works in stochastic dynamics are Foster and Young [20], Kandori, Mailath, and Rob [21], Young [22], and Binmore, Samuelson, and Vaughan [23]. A comprehensive presentation is offered in Young [24].

At first, this paper describes an agent-driven, stochastic evolutionary model featuring heuristic learning. In each new period, players observe whether their past choices have been best replies to those of their opponents. If not, then "disappointment" emerges, and the players present a tendency to switch to alternative strategies in the future. Individual choice based on the avoidance of disappointment or regret has been discussed in the literature by various authors [25–29]. A comparative presentation of disappointment and regret aversion models appears in Grant *et al.* [30]. Our paper presents a few stylized, novel ways to translate the core idea into an evolutionary game theoretic context.

In Loomes and Sugden seminal contribution [25], agents experience regret when some choice they made proves to be less successful (in payoff terms) than another option they did not choose; individuals are aware of this effect, and make their choices so as to avoid regret. In our model, individuals, being

[1] For comprehensive reviews, see [7–8]. More recent texts include [9–10].

boundedly rational, do not take proactive measures to circumvent regret, but rather, they act upon realized disappointing outcomes. The revision protocols that we use to implement this idea are close in character to the adaptive procedure studied in Hart and Mas-Colell [31], where the concept of '*regret matching*' is introduced. While regret matching requires quite demanding computational abilities from the part of the agents, the model presented here is more heuristic and deals with less sophisticated players (not necessarily in the sense that they are less smart, but mainly because they have shorter memory and lack perfect knowledge of their surroundings).

The paper explores these dynamics by use of simulation software, confining the analysis to 2 × 2 symmetric games. The aim is to gain insights on the properties of the different revision protocols (each one of which corresponds to a specific behavioral type), and, thus, investigate the possibility that some of these types consistently outperform the rest. The software also allows for stochastic shocks in the form of 'intruders' (i.e., preprogrammed automata), who may be matched, with positive probability, with members of the original population. Allowing for such perturbations is important, because they account for random errors or even calculated deviations (or interventions), which can potentially affect the evolutionary course, often in unexpected ways. Simulation software programs are, in any case, being used increasingly in the literature in the study of evolutionary games; see, for instance, [32–35].

## 3. The Model

### *3.1. Theoretical Background*

Let $I = \{1, 2,\ldots, N\}$ be the set of players, and for each player $i$, let $S_i$ be her finite set of pure strategies. Following Weibull [36], each player's pure strategies are labeled by positive integers. Thus, $S_i = \{1, 2,\ldots, m_i\}$, for some integer $m_i > 1$. The set of pure strategies in the game, denoted $S$, is the Cartesian product of the players' pure strategy sets. For any strategy profile $s\in S$ and player $i\in I$, let $\pi_i(s)\in R$ be the associated payoff to player $i$. Let $\pi{:}S\rightarrow R^n$ be the combined pure-strategy payoff function of the game (that is, the function assigning to each pure-strategy profile the vector of payoffs $(\pi_1(s), \pi_2(s),\ldots, \pi_n(s))$. With the above notation, the game is summarized by the triplet $G = (I, S, \pi)$.

We use $s^{-i}$ to denote a pure strategy combination of all players except $i$ and define the set of $i$'s pure best replies to a strategy combination $s^{-i}$ to be the nonempty finite set $BR_i = \{h\in S_i : \pi_i(h, s^{-i}) \geq \pi_i(k, s^{-i})\ \forall\ k\in S_i\}$. We define the set of $i$'s pure worst replies to a strategy combination $s^{-i}$ to be the nonempty finite set $WR_i = \{w\in S_i: \pi_i(w, s^{-i}) \leq \pi_i(k, s^{-i})\ \forall\ k\in S_i\}$. We define the *maximum disappointment* $\Delta_i\in R$ as $\Delta_i = \max\{\pi_i(h, s^{-i}) - \pi_i(w, s^{-i})\}$, for all $s^{-i}$, $h\in BR_i$, $w\in WR_i$. In other words, for each possible $s^{-i}$, we calculate the difference between the payoffs associated with the best reply and the payoffs associated with the worst reply to $s^{-i}$; thus, we have $m_1m_2\ldots m_{i-1}m_{i+1}\ldots m_n$ non-negative real numbers. $\Delta$ is the maximum of these numbers.

Suppose that $G$ is played repeatedly in discrete time periods $t = 0, 1, 2,\ldots, T$. We denote $\pi_{t,i}(c_t, s_t^{-i})$ player $i$'s payoffs at time $t$, when they choose $c\in S_i$ (denoted $c_t$), assuming the others chose $s^{-i}$ (denoted $s_t^{-i}$). Similarly, we use $BR_{t,i}$ to denote the set of best replies of player $i$ at time $t$, when the opponents choose $s_t^{-i}$. The following subsections describe three revision protocols, by providing different possible probability distributions used by $i$ at time $t$+1, contingent on $i$'s behavioral traits.

### *3.2. The 'Short-Sightedness' Protocol*

The short-sightedness protocol is described by the following probability distribution:

$$p_{t+1,i}(y) = \mu\cdot(\pi_{t,i}(h_t, s_t^{-i}) - \pi_{t,I}(c_t, s_t^{-i}))/(m_i - 1)\cdot\Delta, \text{ for all } y \neq c,\ y\in S_i,$$

$$p_{t+1,i}(c) = 1-\mu\cdot(\pi_{t,i}(h_t, s_t^{-i}) - \pi_{t,i}(c_t, s_t^{-i}))/\Delta,$$

$$\text{where } h_t\in BR_{t,i},\ 0 < \mu \leq 1. \tag{1}$$

In words, unless player $i$ chooses a best reply at time $t$, she may switch to any one among her alternative strategies with equal probability, one that is proportional to the ratio of the payoffs foregone at time $t$ (the 'disappointment' at time $t$) over the maximum payoffs that could have been lost (the 'maximum disappointment' $\Delta$). Thus, the greater the loss of payoffs at $t$ (due to one's failure of selecting a best reply), the higher the probability that $i$ switches to some alternative strategy. Values of $\mu$ less than 1 indicate inertia, which increases as $\mu$ decreases[2]. Agents conforming to (1) are clearly short-sighted, for they just take into account what has happened in the previous round only.

### *3.3. The 'n-Period Memory' Protocol*

In this revision protocol, players are depicted as more rigid or patient, as they switch only after experiencing $n$ disappointing outcomes in a row (e.g., in rounds $t$, $t$–1, … $t$–$n$+1). The probability distribution of this protocol is given by (2) below:

$$\text{If } \Pi_{\tau=t-n+1}^{t}(\pi_{\tau,I}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i})) \neq 0,$$
$$\text{then } p_{t+1,i}(y) = \Sigma_{\tau=t-n+1}^{t}(\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i}))/((m_i - 1)n\Delta),$$
$$\text{for all } y \neq c,\ y\in S_i,\ p_{t+1,i}(c) = 1 - \Sigma_{\tau=t-n+1}^{t}(\pi_{\tau,i}(h_\tau, s_\tau^{-i})-\pi_{\tau,i}(c_\tau, s_\tau^{-i}))/n\Delta. \tag{2}$$
$$\text{If } \Pi_{\tau=t-n+1}^{t}(\pi_{\tau,I}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i})) = 0, \text{ then } p_{t+1,i}(y) = 0,$$
$$\text{for all } y \neq c,\ y\in S_i,\ p_{t+1,i}(c) = 1, \text{ where } h_\tau\in BR_{\tau,i}, \text{ and } \pi_{\tau,i}(h_\tau, s_\tau^{-i}) = \pi_{\tau,i}(c_\tau, s_\tau^{-i}) = 0 \text{ when } \tau<0.$$

If $n$ = 1, (2) collapses to (1) for the special case where $\mu$ = 1; if $n$ > 1, inertia comes into play with $n$ reflecting the agent's 'resistance' to switching following a string of disappointing outcomes. Note that protocol (2) implies that all past $n$ rounds matter the same. This is plausible to the extent that $n$ is quite small, and that agent $i$ gets involved in the same interaction frequently enough. Discounting of older disappointing outcomes will be considered below.

---

[2] As agent $i$ is assumed as capable of acknowledging when they could have earned more payoffs, it would probably be more realistic to argue that, if $i$ is to change their current strategy, then they are not going to select another strategy at random (as implied by the above distribution), but they shall choose a strategy belonging to the set of best replies to the opponents' choice at $t$ (i.e., the set $BR_{t,i}$). In the case of 2 × 2 games (that shall be studied here), it is obvious that this issue is not a concern. Obviously, for larger games, this protocol reflects a very weak form of learning, and should probably be modified in accordance to the sophistication one would wish to endow the individuals with.

### *3.4. The 'Additive Disappointment' Protocol*

Let $v_i \in R^{T+1}$, $v_i = [v_{0,i}\ v_{1,i} \ldots v_{T,i}]$ and define $v_i$ as: $v_{0,i} = 1$; $v_{\tau,i} = v_{\tau-1,I} + 1$, $\tau \geq 1$, if $c_\tau = c_{\tau-1}$; $v_{\tau,i} = 1$, $\tau \geq 1$, if $c_\tau \neq c_{\tau-1}$. This vector effectively keeps track of when player $i$ changes their strategy, and of how many rounds each strategy has lasted. For example, if player $i$ chooses strategy $\beta \in S_i$ at $t = 0$, $t = 1$, $t = 2$, then strategy $\gamma \in S_i$, $\gamma \neq \beta$ at $t = 3$, and then strategy $\delta \in S_i$, $\delta \neq \gamma$ at $t = 4$ and $t = 5$, then the first six elements of $v_i$ will be 1, 2, 3, 1, 1, 2. Hence, whenever we see that $v_{\tau,i}$ is equal to 1, we can know that a change of strategy happened at $t = \tau$, and that $i$ had been choosing their previous strategy (the one they played until $t = \tau-1$) for a number of rounds equal to $v_{\tau-1,i}$. We can now define the *additive disappointment* protocol by (3) below:

$$\begin{gathered}\text{If } \Sigma_{\tau=t-v_{t,i}+1}^{\ t} \rho^{t-\tau} \cdot (\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-I})) \geq Z, \text{ then } p_{t+1,i}(y) = 1/(m_i - 1), \\ \text{for all } y \neq c,\ y \in S_i,\ p_{t+1,i}(c) = 0. \\ \text{If } \Sigma_{\tau=t-v_{t,i}+1}^{\ t} \rho^{t-\tau} \cdot (\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-I})) < Z, \text{ then } p_{t+1,i}(y) = 0, \\ \text{for all } y \neq c,\ y \in S_i,\ p_{t+1,i}(c) = 1, \text{ where } h_\tau \in BR_{\tau,i},\ 0 < \rho \leq 1, \text{ and } Z \in R^+.\end{gathered} \tag{3}$$

The above reflects players who change their current strategy $c$ at $t$+1 if and only if the amassed disappointment from the previous $v_{t,i}$ rounds (where $c$ was played) surpasses a predefined threshold $Z$. This protocol describes players who abide with a strategy for a number of rounds until total disappointment exceeds some subjective "tolerance" level. In each new period, the disappointment received from the previous $v_{t,i}$ rounds is discounted at rate $\rho$. For simplicity, the population will be thought of as homogeneous in terms of parameters $\rho$ and Z.

## 4. The Simulation

The software assumes a population of 1,000 agents and is amenable to symmetric 2 × 2 games[3]. The user specifies a number of iterations $x$. In each iteration or period, agents are randomly paired. The user may introduce a positive probability $a$ with which an agent is, unbeknownst to her, matched against an "intruder"; that is, an 'outsider' who does not belong to the original population and whose behavior has not evolved endogenously. To keep things simple, it is assumed that intruders act like automata, selecting their first strategy with probability $b$ in every round and independently of their own past experience. In effect, they represent stochastic perturbations of the original population's evolutionary process. Any values for $a$ and $b$ in [0,1] are permitted (with 0.1 and 0.5 respectively being the default values). After the end of the simulation, each player will have participated in approximately $x \cdot (2 - a)/1{,}000$ rounds. With no intruders ($a = 0$), this number collapses to $x/500$.

At the outset, agents choose strategies with predefined (by the user) probabilities, with $p_t$ denoting the fraction of the population that chooses the first strategy at time $t$. As we focus on symmetric games, this is also the probability that a random individual chooses their first strategy at $t$. The default value for $p_0$ is 0.5.

---

[3] The software was written by the author in Microsoft Visual Basic 6.0. Random numbers are generated by use of the language's Rnd and Randomize functions.

One crucial difference of our model with the canonical version of evolutionary game theory lies in that it makes no allusion to expected values, but rather, uses realized ones. This leads to a more plausible representation of the evolutionary course, consistent with the empirical assertion that someone's current choices will reflect their tendency to avoid unpleasant past experiences due to perceived erroneous choices. The specific mechanics of the tendency to switch strategies following such 'disappointment' depends on the three behavioral types described in the previous section. The modifications of revision protocols (1), (2), and (3) necessary to introduce them into the software presented therein are straightforward. As only two players are randomly selected to participate in each round, the probability distributions (1), (2), and (3) are valid only for the periods where player $i$ participates. Thus, periods $t = 0, 1, 2,\ldots, T$, as used in these distributions, are no longer all the rounds of the game, but only those involving player $i$[4].

The following subsections present the modifications to protocols (1), (2), and (3), as implemented by the software program.

### *4.1 Short-Sightedness*

We denote the currently used strategy with $c$, and the alternative one with $y$, while the opponent's current strategy is $s^{-i}$. The time periods indicate the rounds where $i$ is randomly selected to play the game, while parameter $\mu$ (which determines the maximum probability of a switch) has been hard-coded equal to 1/3. Protocol (1) is rewritten as follows:

$$p_{t+1,i}(y) = (\pi_{t,i}(h_t, s_t^{-i}) - \pi_{t,i}(c_t, s_t^{-i}))/3\Delta,\ y \neq c,\ y \in S_i,$$
$$p_{t+1,i}(c) = 1 - (\pi_{t,i}(h_t, s_t^{-i}) - \pi_{t,i}(c_t, s_t^{-i}))/3\Delta,$$
$$\text{where } h_t \in BR_{t,i}. \tag{1'}$$

### *4.2 Three-Period Memory*

In addition to its intuitive appeal, the introduction of behavioral inertia helps explain why the evolutionary process manages to escape its original state $p_0$ (see Subsection 4.5 for an example). Probability distribution (2') below is a special case of revision protocol (2) for 2 × 2 games and with $n = 3$. Here, switching to the other strategy is a possibility only when the current strategy has generated three disappointing outcomes in a row; an arbitrary, yet quite plausible assumption (supported also by popular phrases such as 'to err thrice…').

$$\text{If } \Pi_{\tau=t-2}^{t}(\pi_{\tau,i}(h_\tau, s_\tau^{-I}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i})) \neq 0,$$
$$\text{then } p_{t+1,i}(y) = \Sigma_{\tau=t-2}^{t}(\pi_{\tau,i}(h_\tau, s_\tau^{-I}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i}))/3\Delta,\ y \neq c,\ y \in S_i,$$
$$p_{t+1,i}(c) = 1 - \Sigma_{\tau=t-2}^{t}(\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i}))/3\Delta. \tag{2'}$$

$$\text{If } \Pi_{\tau=t-2}^{t}(\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i})) = 0 \text{ then } p_{t+1,i}(y) = 0,\ y \neq c,\ y \in S_i,\ p_{t+1,i}(c) = 1,$$
$$\text{where } h_\tau \in BR_{\tau,i}, \text{ and } \pi_{\tau,i}(h_\tau, s_\tau^{-i}) = \pi_{\tau,i}(c_\tau, s_\tau^{-i}) = 0 \text{ when } \tau < 0.$$

---

[4] As this adjustment causes no ambiguity, we will simplify notation by not adding $i$ subscripts to the time periods, as a more formal representation would require.

Three-period memory implies patient players who do not switch strategies at the first, or second setback. Change of current strategy $c$ will not be considered if $c$ has been the best reply in any one of the last three periods.

### *4.3 Additive Disappointment*

Additive disappointment offers a less stylized variation of the previous protocol: agent $i$'s memory extends back to as many periods as the number of rounds that $i$ has played $c$. The software assumes that disappointment in round $t–\tau$ is discounted at a $0.9^{\tau}$ rate. The threshold value $Z$ is fixed equal to two times the maximum disappointment level $\Delta$. Hence, and given the discounting, $i$ cannot switch to $y \neq c$ in fewer than three consecutive disappointing rounds.

$$\begin{gathered}\text{If } \Sigma_{\tau=t-v_{t,i}+1}^{\ t} 0.9^{t-\tau}\cdot(\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i})) \geq 2\Delta, \\ \text{then } p_{t+1,i}(y) = 1,\ y \neq c,\ y \in S_i,\ p_{t+1,i}(c) = 0. \\ \text{If } \Sigma_{\tau=t-v_{t,i}+1}^{\ t} 0.9^{t-\tau}\cdot(\pi_{\tau,i}(h_\tau, s_\tau^{-i}) - \pi_{\tau,i}(c_\tau, s_\tau^{-i})) < 2\Delta, \\ \text{then } p_{t+1,i}(y) = 0,\ y \neq c,\ y \in S_i,\ p_{t+1,i}(c) = 1, \\ \text{where } h_\tau \in BR_{\tau,i}.\end{gathered} \tag{3'}$$

Under this protocol, players switch strategies when they feel they have 'had enough' of the disappointment due to their strategy choice. The specified discount factor and disappointment threshold, while arbitrary, models, nicely, agents who adopt satisficing behavior in that they stick to some possibly suboptimal strategy until their tolerance level (factoring in the relative importance of more recent disappointment) is exceeded.

### *4.4 The Games*

Table 1 shows the reference games to be used in the simulation:

**Table 1.** Five classic games.

| | 1 | 2 |
|---|---|---|
| 1 | 3,3 | 0,5 |
| 2 | 5,0 | 1,1 |

Game 1: Prisoner's Dilemma

| | 1 | 2 |
|---|---|---|
| 1 | 2,2 | 0,0 |
| 2 | 0,0 | 2,2 |

Game 2: Coordination

| | 1 | 2 |
|---|---|---|
| 1 | 2,2 | 0,0 |
| 2 | 0,0 | 1,1 |

Game 3: Hi-Lo

| | 1 | 2 |
|---|---|---|
| 1 | -2,-2 | 2,0 |
| 2 | 0,2 | 1,1 |

Game 4: Hawk-Dove

| | 1 | 2 |
|---|---|---|
| 1 | 3,3 | 0,2 |
| 2 | 2,0 | 1,1 |

Game 5: Stag-Hunt

In each of these games, the "social welfare" standpoint rejects certain outcomes (as inferior). Unfortunately, individually rational action often leads players to these very outcomes. In the '*Prisoner's Dilemma*' and '*Hawk-Dove*' games, for example, the socially desirable strategies cooperation and dovish behavior are trumped by defection and hawkish behavior (as a part of a mixed strategy) respectively, while, in '*Coordination*', rationality alone cannot prevent coordination failure—and the same applies for '*Hi-Lo*' and '*Stag-Hunt*'. Our interest is to see whether behavior evolving according to the above protocols leads to results that differ substantially from those suggested by standard game theory. The simulation results are described below.

### *4.5 Simulation Results: Case without Random Perturbations*

In this subsection, we assume that there are no intruders ($a$ = 0), and therefore, members of the population always interact between themselves. Interactions with intruders shall be seen as random 'shocks', and they will be presented in subsection 4.6.

#### 4.5.1. Prisoner's Dilemma and Hawk-Dove

Unsurprisingly, in the '*Prisoner's Dilemma'* defection dominates the entire population, no matter the behavioral type and the initial conditions. As in standard evolutionary game theory, defection always yields zero disappointment, while cooperation always yields positive disappointment; thus, the agents shall eventually switch to defection, regardless of their behavioral code, insofar this disappointment is assumed to generate a tendency to defect, if one cooperates. Of course, things might have turned out otherwise if disappointment were to be defined differently. Here, disappointment is the feeling one gets when comparing one's actual payoff to what it would have been had one chosen a best reply strategy, the opponent's choice being a given. Naturally, if disappointment were to increase in proportion to one' share of foregone collective payoffs, then mutual cooperation would be a possibility, as long as players cared about the welfare of both participants as a group, and not just for their own performance in the game. One way to incorporate this consideration would typically be to change the payoffs of the game in order to reflect the increase in utility possibly derived from a mutually beneficial outcome and then rerun the simulation for the amended game. This would commonly give us a game with the strategic structure of '*Stag-Hunt*', the simulation results for which are presented below.

In '*Hawk-Dove*', short-sightedness causes evolution to converge to a state where around 41% choose the hawkish strategy *for any initial condition*; a level of aggression considerably higher to that predicted by standard evolutionary game theory, where the evolutionary stable equilibrium is $p$ = 1/3. The aggression, in fact, grows further under the three-period memory protocol to, approximately, $p$ = 0.46. The explanation for this last result is that the players' relative rigidity acts as an enabler for the existence of more 'hawks' in the population; under "three-period memory", players who behave aggressively need to be paired with other aggressive players in three consecutive rounds before changing their strategy to 'dove', and this persistence tolerates more 'hawks' in the aggregate than in the case of myopic players who switch to the other strategy more readily. Finally, under additive disappointment, the system converges to approximately $p$ = 0.41 (same level approximately as under short-sightedness). This rate of aggression decreases dramatically as the payoff consequences of a (Hawk–Hawk) outcome become more disastrous: if, for example, the payoffs of the (Hawk–Hawk) outcome are changed from (–2,–2) to (–4,–4), then the new equilibrium becomes $p$ = 0.25 (N.B., the other protocols give $p$ = 0.33 (short-sightedness) and $p$ = 0.43 (three-period memory)). This decrease in 'hawks' is explained on the grounds that, as the disappointment caused by a conflictual outcome became four times greater than the disappointment generated by mutual acquiescence (Dove–Dove) on the occasion that 'hawks' cross paths, it is much more probable that the threshold value determining whether a change of strategy will happen or not is surpassed (as opposed to the case where two 'doves' meet).

4.5.2. Coordination, Hi-Lo, and Stag Hunt

In these three variants of the coordination problem, standard evolutionary game theory admits each game's pure Nash equilibria as evolutionarily stable. Which of the two obtains depends on the initial conditions: In '*Coordination*' and '*Stag Hunt*', the top left (bottom right) equilibrium, $p = 1$ ($p = 0$), will emerge if, at the outset, more (less) than 50% of the population chose the first strategy (i.e., if $p_0 > 1/2$). In '*Hi-Lo*', the top left equilibrium requires a smaller initial critical mass in order to dominate ($p_0 > 1/3$).

In contrast, under the *short-sightedness* protocol presented in section 4.1, above, convergence to one of the two equilibria is not as straightforward. In the pure '*Coordination*' game, while the system has no tendency to leave these equilibria once there (as no individual ever experiences disappointment at these equilibria), for any other initial condition ($0 < p_0 < 1$) the system will perform a random walk, oscillating back and forth in each round with equal probabilities, to only rest if one the two equilibria is accidentally reached. The explanation for this lies in that both players have the same probability of changing their strategy when coordination fails; therefore, the probability for $p_t$ to increase in some round $t$ is the same as the probability for $p_t$ to decrease (and equal to $2/9 = \mu(1 - \mu)$); and if both players switch their strategies, or if neither does, then $p_t$ remains unchanged.

It follows that under the short-sightedness protocol, agents caught in a pure '*Coordination*' problem do a poor job of achieving coordination, resembling the embarrassing situation of two people trying to avoid collision when walking toward one another in some corridor. This result also shows why *some* inertia may be useful: for if $\mu$ were equal to 1, then, in any instance where one player would choose the first strategy and the other player would choose the second strategy, then both would switch to their other strategy with probability 1, and, hence, $p_t = p_0$ for all $t$.

The situation is different when the short-sightedness scenario is used in the context of '*Hi-Lo*'. In that game, the fact that the equilibria are Pareto-ranked helps the system converge to the optimal equilibrium $p = 1$, for any initial condition. The only exception is, naturally, the extreme case where $p_0 = 0$. In short, the efficient outcome is attained as long as a minor fraction of individuals is opting, initially, for the first strategy, because an instance of non-coordination in this game brings greater disappointment to the player who chose the second strategy; hence, the probability with which the player who chose the second strategy will switch to the first strategy is greater than the probability with which the player who chose strategy 1 will switch to strategy 2. This explains the convergence to the efficient outcome from any state except for the one where (nearly) all players choose the second strategy.

Turning now to the *three-period memory* protocol, in pure '*Coordination*', the results coincide with those of standard evolutionary game theory (if $p_0 = 0.5$, then the system converges to either $p = 0$ or $p = 1$ with equal probabilities; if $p_0 <(>) 0.5$, it tends to $p = 0(1)$). Here, the relative rigidity of the players is what enables them to arrive at an equilibrium, for switching to the other strategy may only happen after three consecutive disappointing outcomes, which makes a change of strategy more probable when fewer people are choosing it. In juxtaposition, applying the three-period memory protocol to '*Hi-Lo*' leads to the result that the necessary initial condition for emergence of the efficient equilibrium is approximately $p_0 > 0.42$. When the equilibria are Pareto-ranked, and unlike the situation in which agents are short-sighted, the evolution of the optimal equilibrium requires a lot more initial

adherents (i.e., a much higher $p_0$). Clearly, the players' relatively high inertia may inhibit the evolution of the Pareto optimal outcome since, if a critical mass of a least 42% of players opting for the Pareto optimal strategy is not present from the outset, it is sufficiently likely that those who choose it will experience three disappointing results in a row, causing them to switch to the strategy that corresponds to the suboptimal equilibrium.

On the other hand, under the *additive disappointment* protocol, coordination at the optimal equilibrium of '*Hi-Lo*' is more probable (though still less so than in the short-sightedness protocol case) as the efficient equilibrium's catchment basin is approximately $p_0 > 0.17$. Interestingly, when the suboptimal strategy (strategy 2) carries less risk in case of coordination failure, as in the '*Stag Hunt*' game, the way in which disappointment affects the players makes little difference: under both the three-period memory and the additive disappointment protocols, efficiency is guaranteed as long as $p_0 > 1/2$, and condemned if $p_0 < 1/2$.

It is worth noticing that the analysis is sensitive to the relative attractiveness of the efficient equilibrium (just as the relative unattractiveness of the conflictual outcome made a crucial difference in '*Hawk-Dove*'). If, for example, we change the payoffs of the efficient outcome of '*Stag-Hunt*' from (3,3) to (4,4), then we see that, under short-sightedness, the system converges to the efficient equilibrium from *any* initial state (except for the case where $p_0 = 0$). However, this can also work the other way round: if we make the inferior outcome less unattractive, then we get the same effect in reverse. For example, if we change the payoffs of the sub-optimal outcome of '*Stag-Hunt*' from (1,1) to (1.5,1.5), then the system shall always converge to $p = 0$ (unless $p_0 = 1$). In this last example, the three-period memory and the additive disappointment protocols lead to the efficient equilibrium as long as $p_0 > 0.55$ and $p_0 > 0.74$, respectively.

Table 3 summarizes the above results for the games of Table 1 and the amended games featured in Table 2.

It is now clear that the agents' behavioral type is a crucial determinant of the evolutionary path. The players' attitude to disappointing resolutions may not make a difference in games with unique dominant strategy equilibria, like the '*Prisoner's Dilemma*', but it does so in games featuring multiple Nash/evolutionary equilibria, e.g., '*Hawk–Dove*', pure '*Coordination*', '*Hi-Lo*', or '*Stag-Hunt*'.

Our first insight is that, while short-sighted players perform poorly when coordinating on equally desirable equilibria (e.g., pure '*Coordination*'), they may be more adept than players with longer memories at sidestepping paths that railroad them toward inefficient equilibria (as in '*Hi-Lo*'). The same may be also true for coordination type games where Pareto-efficiency and an aversion to the worst outcome may pull players in different directions, e.g., '*Stag-Hunt*'. On the other hand, as the relative benefits from the efficient outcome decrease, the agents' myopia may have the opposite effect (recall '*Stag-Hunt* #3').

**Table 2.** Games 4 and 5, amended.

| | 1 | 2 |
|---|---|---|
| 1 | -4,-4 | 2,0 |
| 2 | 0,2 | 1,1 |

Game 6: Hawk-Dove #2

| | 1 | 2 |
|---|---|---|
| 1 | 4,4 | 0,2 |
| 2 | 2,0 | 1,1 |

Game 7: Stag-Hunt #2

| | 1 | 2 |
|---|---|---|
| 1 | 3,3 | 0,2 |
| 2 | 2,0 | 1.5,1.5 |

Game 8: Stag-Hunt #3

**Table 3.** Simulation results in the case without random perturbations.

| | **Short-Sightedness** | **Three-Period Memory** | **Additive Disappointment** |
|---|---|---|---|
| **Prisoner's Dilemma** | | $p = 0$ from any initial state | |
| **Coordination** | Randomness, possible convergence to $p = 0$ or $p = 1$ | $p = 1$, if $p > 1/2$ at $t = 0$<br>$p = 0$, if $p < 1/2$ at $t = 0$ | $p = 1$, if $p > 1/2$ at $t = 0$<br>$p = 0$, if $p < 1/2$ at $t = 0$ |
| **Hi-Lo** | $p = 1$, if $p > 0$ at $t = 0$<br>$p = 0$, if $p = 0$ at $t = 0$ | $p = 1$, if $p > 0.42$ at $t = 0$<br>$p = 0$, if $p < 0.42$ at $t = 0$ | $p = 1$, if $p > 0.17$ at $t = 0$<br>$p = 0$, if $p < 0.17$ at $t = 0$ |
| **Hawk-Dove** | $p \approx 0.41$ | $p \approx 0.46$ | $p \approx 0.41$ |
| **Hawk-Dove #2** | $p \approx 0.33$ | $p \approx 0.43$ | $p \approx 0.25$ |
| **Stag-Hunt** | | Same results as '*Coordination*' | |
| **Stag-Hunt #2** | | Same results as '*Hi-Lo*' | |
| **Stag-Hunt #3** | $p = 0$, if $p < 1$ at $t = 0$<br>$p = 1$, if $p = 1$ at $t = 0$ | $p = 1$, if $p > 0.55$ at $t = 0$<br>$p = 0$, if $p < 0.55$ at $t = 0$ | $p = 1$, if $p > 0.74$ at $t = 0$<br>$p = 0$, if $p < 0.74$ at $t = 0$ |

Agents described by the additive disappointment protocol fare better in antagonistic interactions; their attitude of sticking to a strategy (unless the amassed disappointment from previous rounds surpasses some threshold) is bound to turn them into a peaceful group of people, not necessarily because they favor peace, but because they are contended more easily and have no incentive to strive for more, at peace's expense. Moreover, these agents perform remarkably well in '*Hi-Lo*', '*Stag-Hunt*', and '*Stag-Hunt* #2' (albeit worse than short-sighted agents), but not so well in '*Stag-Hunt* #3', where the catchment area of the efficient equilibrium is relatively small.

The three-period memory protocol credits the agents with some level of sophistication. Their elevated inertia does not seem to be in their favor in several cases, especially in '*Hawk-Dove*', where the resulting aggression is too high, and in '*Hi-Lo*' or '*Stag-Hunt* #2', where the basin of attraction of the efficient equilibrium is smaller than the other behavioral types; however, their sense of caution pays off in '*Stag-Hunt* #3', where they are ultimately driven to the optimal outcome even for relatively low initial $p_0$ values. These players are sometimes not flexible enough to let the evolutionary course work in their favor, but this very rigidity is what may protect them against possibly unpleasant situations (such as being attracted by the sub-optimal equilibrium in '*Stag-Hunt* #3' from any initial state except $p_0 = 1$, as happens under short-sightedness).

### *4.6 Simulation Results: Case with Random Perturbations*

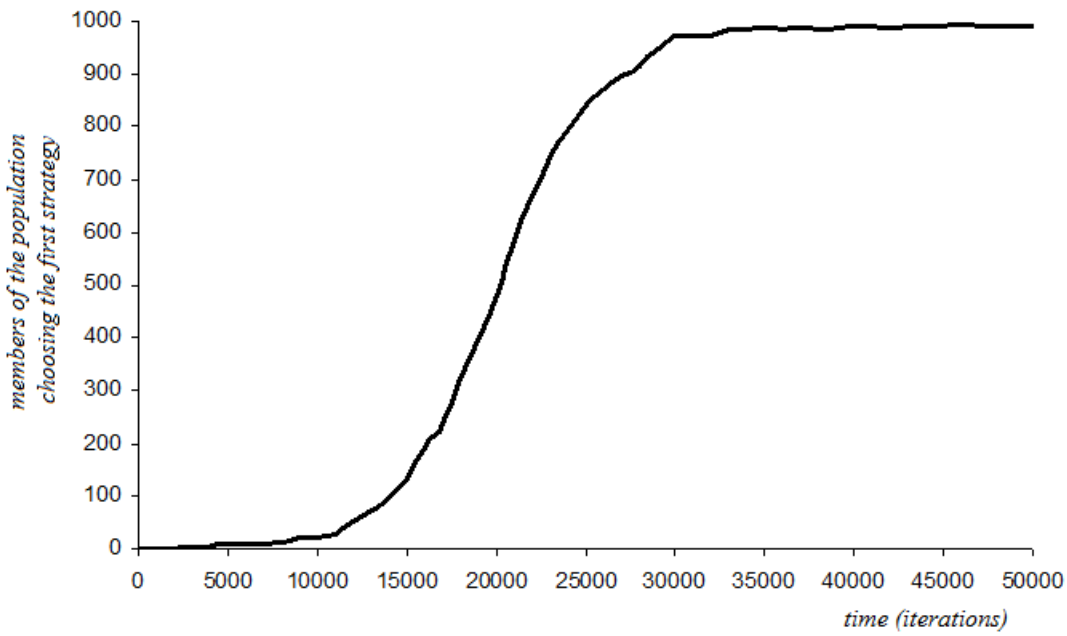
#### 4.6.1. Short-Sightedness with Stochastic Perturbations

We have already noticed how short-sighted players may be heavily influenced by minor perturbations. To explore this further, the simulation software has been augmented potentially to include 'intruders'. The latter interact with our population members with probability $a$ (in each iteration) and choose the first strategy with probability $b$. Their presence may be interpreted either as a random error or a deliberate intervention—perhaps from a third party aiming to help the original population arrive at a desired equilibrium. These perturbations might as well be considered as noise that

sometimes enables (as shall be seen below) the dynamics to leave a catchment area and enter another basin of attraction.

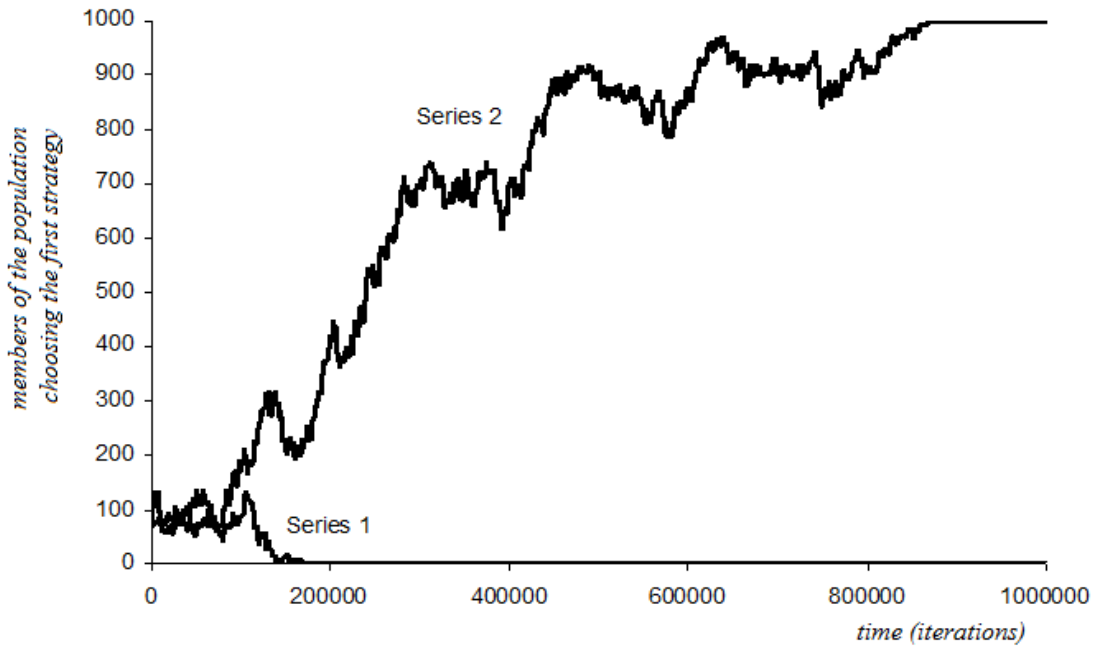
To illustrate, let us consider '*Hi-Lo*' under short-sightedness. In the case without perturbations, we saw that the efficient equilibrium is threatened only if the whole population is stuck, at the very beginning, in the suboptimal equilibrium. Naturally, the introduction of only a few intruders is enough to guarantee convergence to the efficient equilibrium $p = 1$. Figure 1 demonstrates this under the assumption that $p_0 = 0$, $a = 0.01$, and $b = 0.1$: as time goes by (horizontal axis), the number of individuals (out of 1,000) who choose strategy 1 grows inexorably (the vertical axis depicts the number of individuals who choose their first strategy or, equivalently, $p_t$ multiplied by 1,000). Convergence to the efficient outcome took, in this simulation, around 80 games per person (less than 40,000 iterations in total).

**Figure 1.** '*Hi-Lo*' when all players are initially 'stuck' in the inefficient outcome. The introduction of few intruders sets them on a course to the efficient outcome.




We now turn to '*Stag-Hunt*', where the evolutionary path may take the population to one of the two available equilibria. Without stochastic perturbations, short-sightedness threatened to put the population in an endless drift (see previous section). Typically, if $p_0 = 0.5$, our (non-stochastic) simulation took more than one million iterations for the system to hit one of the two absorbing barriers. Naturally, the closer $p_t$ is to one of the two barriers/equilibria, the more probable convergence is to that equilibrium point. However, the addition of intruders can change this. Consider the case where $p_0 = 0.1$. In Figure 2, Series 1 shows the results of a simulation without intruders: predictably, the proximity of the system's initial condition to the inefficient outcome causes the population to converge toward it quite quickly. However, the addition of a small number of intruders (1% of the population, i.e., $a = 0.01$, who always play the first strategy, i.e., $b = 1$) gives rise to Series 2 and, thus, to a drastically different path. The intruders' presence becomes the catalyst which creates what could be called as 'optimism' within the group, and ultimately drives it towards the optimal equilibrium.

**Figure 2.** '*Stag-Hunt*' when 90% of players are, initially, drawn to the inefficient outcome (Series 1). The introduction of few intruders sets them on a course to the efficient outcome (Series 2).




Notwithstanding the obvious merits of short-sightedness, the implied impatience of the agents and the ease with which they switch to the other strategy may not always be a virtue. Figure 3 demonstrates the point: Here, 80% of the population is drawn to the efficient outcome and yet the presence of a similar number of intruders (as in Figure 2; namely $a = 0.01$ and $b = 0$) causes the evolutionary path to take the population straight into the arms of the inefficient outcome.

Our analysis in the case of no intruders in the previous section suggested that, in games of the '*Stag-Hunt*' structure, short-sighted players may be easy to manipulate, depending on the relative gains from achieving a Pareto superior outcome. Under shortsightedness, players are assumed to experience equal disappointment from an instance of non-coordination in '*Stag-Hunt*', regardless of whether they chose the first of the second strategy. In '*Stag-Hunt* #3', however, the player who chooses the first strategy receives more disappointment than the player who chooses the second strategy, and hence, the intervention for emergence of the efficient equilibrium needs to be more drastic. Figure 4 presents a relevant simulation, with the initial condition $p_0 = 0.5$. Without intruders, the evolutionary course would have taken the population to the sub-optimal outcome (Series 1). However, a sizeable population of intruders may avert this: with $a = 0.5$ and $b = 1$ (that is, if all agents have a 1 in 2 chance of meeting an intruder who always selects the first strategy), the efficient outcome is guaranteed (Series 2). We notice that, on the one hand, the efficient equilibrium is attained, but, on the other hand, we can no longer speak of a minor perturbation or an uncalculated error: the intervention here has to be quite radical.

**Figure 3.** '*Stag-Hunt*' when 80% of players are, initially, drawn to the efficient outcome. The introduction of few intruders sets them on a course to the inefficient outcome.

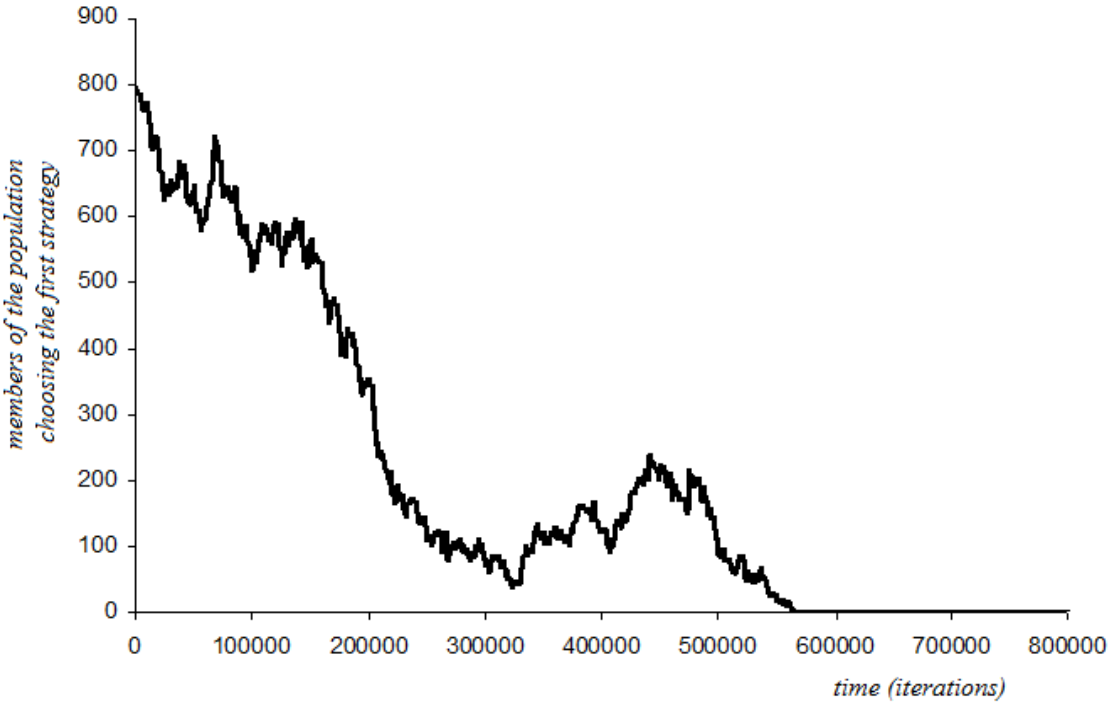
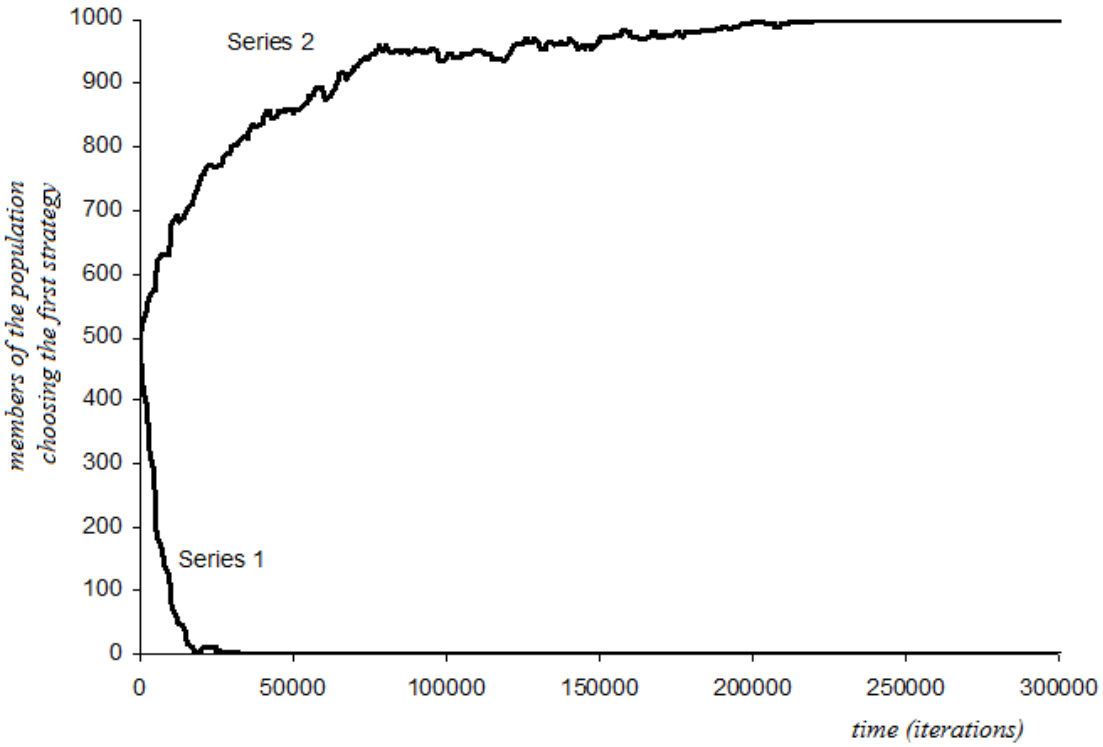



**Figure 4.** '*Stag-Hunt* #3' when 50% of players are, initially, drawn to the efficient outcome. When 50% of them are intruders the efficient outcome is guaranteed (Series 2).



4.6.2. Three-Period Memory with Stochastic Perturbations

The three-period memory behavioral code is generally resistant to shocks. While a minor shock may have dramatic effects under short-sightedness, the same does not apply when agents are more patient, even when the perturbation is far from discrete.

In '*Hawk-Dove*', we found that a three-period memory protocol with no random perturbations increased the players' observed aggression. When a significant probability of meeting a hawkish intruder is introduced, one might be excused to expect a considerable drop in aggression. But that is not what we find in our simulation results. Figure 5 compares a simulation when there are no intruders (Series 1) with one in which there is a 33% probability ($a$ = 0.33) of meeting an intruder who always chooses 'Hawk' ($b$ = 1). The initial condition is $p_0$ = 0.5. We find that, while the percentage of aggressive players indeed decreases, the effect is rather minor (the difference of the two series is less than 10%, which is insignificant *viz.* a perturbation involving 1 in 3 games played by every person).

Turning to '*Hi-Lo*', and recalling that the three-period memory protocol led to suboptimal results for the population as a whole, an infusion of intruders may make the necessary difference as long as their number is high enough. To give one example, in Figure 6 we set $p_0$ = 0.35. Without intruders, as we saw in previous sections, the system will rest at the inefficient equilibrium ($p$ = 0), in contrast to the short-sightedness and the additive disappointment protocols, where $p$ = 1 is the equilibrium. Nothing changes here when the proportion of intruders is small. Series 1 of Figure 6 demonstrates this amply. However, when the proportion of intruders rises to approximately 15%, a different path becomes possible, one that take the population to the optimal outcome. More precisely, for 10 different simulations of the same scenario with $p_0$ = 0.35, $a$ = 0.15, and $b$ = 1, there were six instances of convergence to the efficient equilibrium. Series 2 shows one of these instances (N.B., the smaller $p_0$, the greater the value of $a$ necessary for the system to converge to $p$ = 1).

**Figure 5.** '*Hawk-Dove*' with $p_0$ = 0.5. Series 1: $a$ = 0, Series 2: $a$ = 0.33, $b$ = 1.

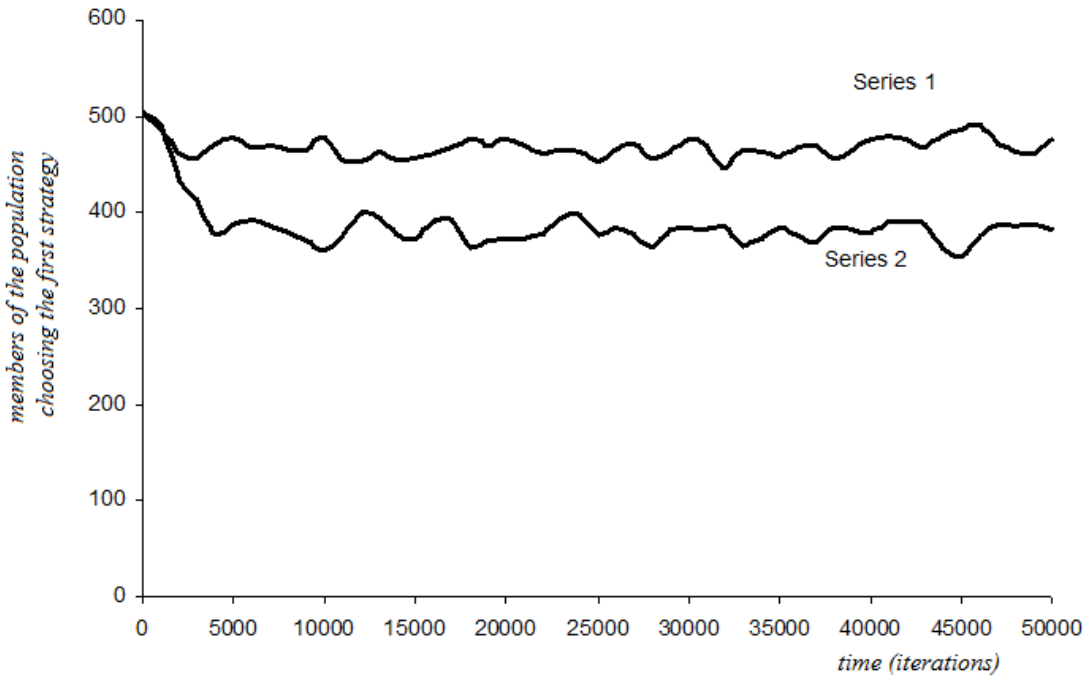
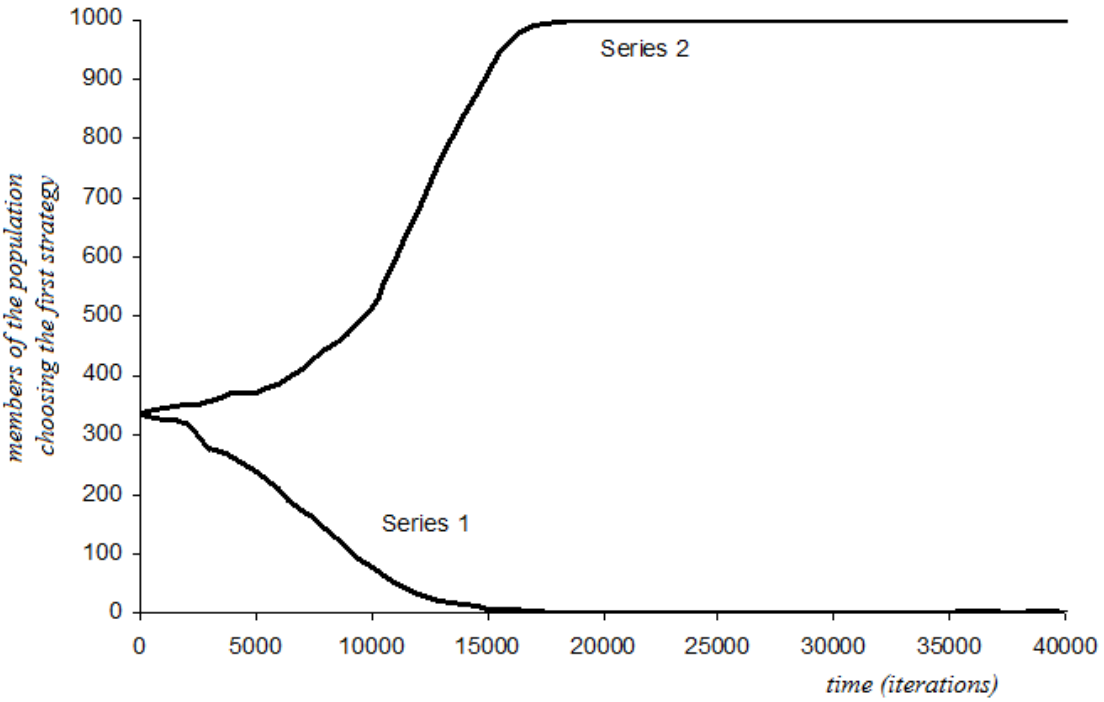


**Figure 6.** '*Hi-Lo*' with $p_0$ = 0.35. Series 1: $a$ = 0.1, $b$ = 1. Series 2: $a$ = 0.15, $b$ = 1.




Meanwhile, in '*Stag-Hunt*', the inefficient result will not be avoided even in the presence of a sizeable population of intruders. Figure 7 shows that even if, say, 40% of the population are drawn initially to the 'good' strategy *and* there is a probability of 15% of meeting with an intruder who also plays the 'good' strategy, the efficient equilibrium (that requires players to choose their 'good' strategies) will not eventuate. Naturally, in some circumstances, such rigidity may turn out to be in the players' favor. In '*Stag-Hunt* #3', with no intruders, we have already observed how short-sighted agents are attracted to the sub-optimal equilibrium (as a player who chooses the second strategy has a smaller probability of changing their strategy than a player who chooses the first strategy when the outcome of some round is (1,2) or (2,1)). In the case of the three-period memory protocol, the sub-optimal outcome has $p_0$<0.55 as its catchment area, and when $p_0$ is slightly greater than that, the efficient equilibrium is not threatened, not even when there is a 10% percent probability of meeting an intruder who always selects the second strategy ($a$ = 0.1, $b$ = 0). Figure 8 offers a relevant simulation with $p_0$ = 0.6. Series 1 emerges when there are no intruders, while Series 2 illustrates the scenario $a$ = 0.1, $b$ = 0. Note how the efficient equilibrium is reached either way, albeit at different speeds depending on the preponderance of intruders.

**Figure 7.** '*Stag-Hunt*' with $p_0 = 0.4$, $a = 0.15$, $b = 1$.

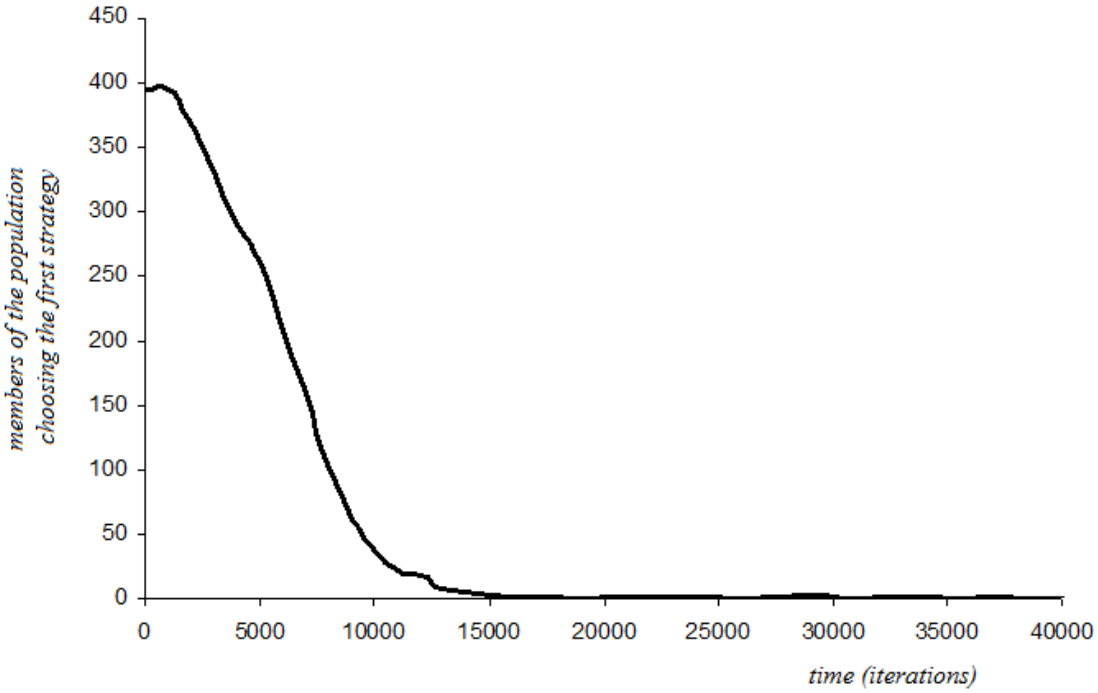
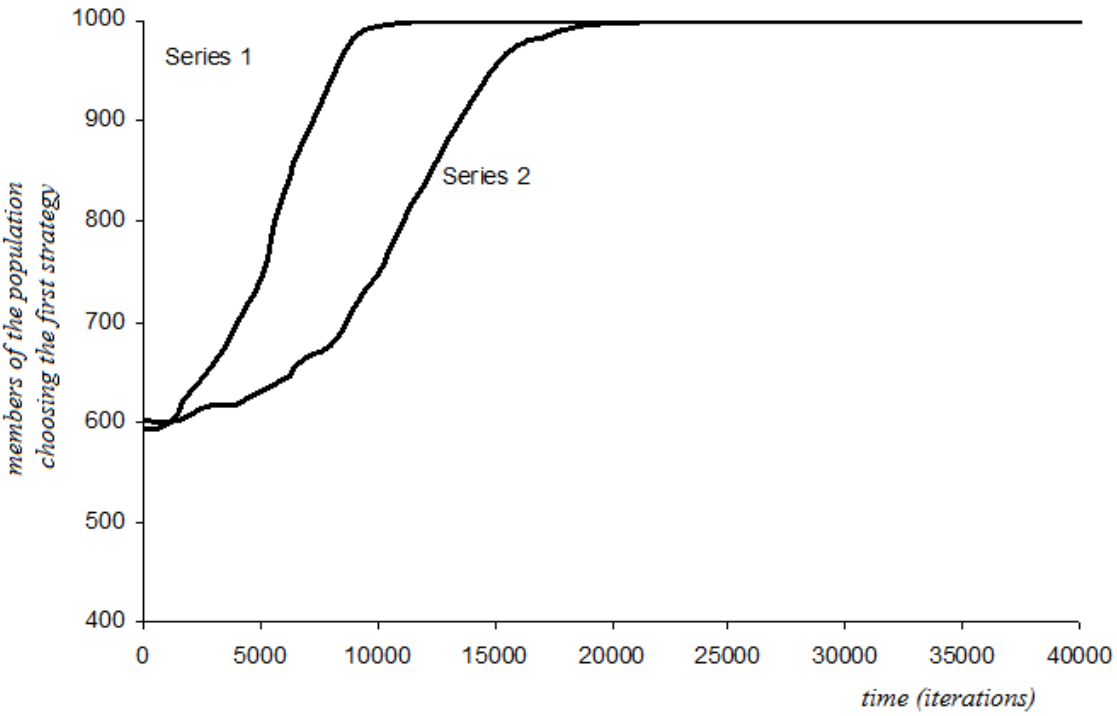



**Figure 8.** '*Stag-Hunt* #3' with $p_0 = 0.6$. Series 1: $a = 0$. Series 2: $a = 0.1$, $b = 0$.

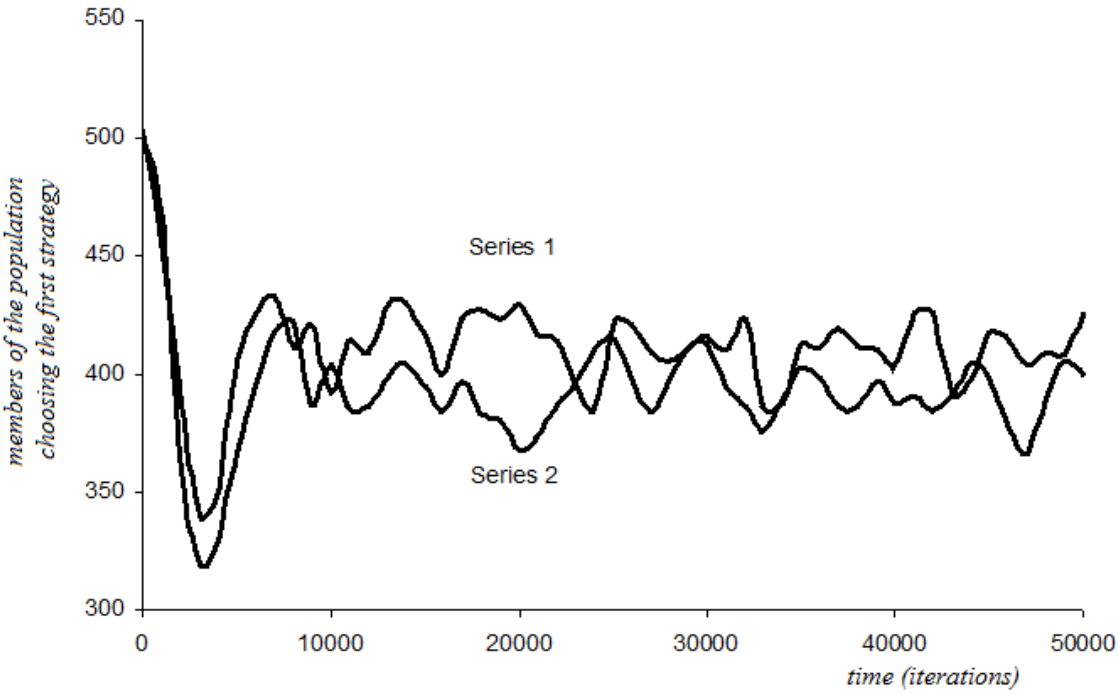


4.6.3. Additive Disappointment with Stochastic Perturbations

Our additive disappointment protocol stands as some kind of middle ground between the impatience of the short-sighted players and the inertia of agents behaving under the three-period memory protocol. This section concludes with several instructive scenarios based on the disappointment protocols.

Under additive disappointment, Figure 9 shows that the presence of intruders lowers the population's aggression rate in '*Hawk-Dove*', although the effect is not distinctly large. In '*Hi-Lo*', the efficient outcome seems to have a surprisingly big basin of attraction ($p_0 > 0.17$). Even if $p_0 < 0.17$, a small perturbation is enough to avert convergence to the sub-optimal equilibrium (see Figure 10 which suggests that players who conform to the additive disappointment protocol are more prone to external influences than agents acting upon three-period memory, albeit they are not as impulsive as the short-sighted players). In '*Stag-Hunt*', we notice a similar effect: Figure 11 shows that lack of intruders means convergence to the socially lesser equilibrium (Series 1, with $p_0 = 0.1$ and $a = 0$), whereas an infusion of 10% intruders suffices to energize a path like that of Series 2. Once more, we find that the population needs only a mild external influence to avoid unpleasant consequences but shows less flexibility when compared to short-sighted players.

**Figure 9.** 'Hawk-Dove' with $p_0 = 0.5$. Series 1: $a = 0$, Series 2: $a = 0.1$, $b = 1$.




The relative inertia of the additive disappointment protocol is also illustrated in Figure 12 which describes the evolutionary course for '*Stag-Hunt* #3' with $p_0 = 0.5$, $a = 0.25$, and $b = 1$. Even though the proportion of intruders is quite large (25%), this is not enough to favor the optimal equilibrium. However, by the same token, the population does not converge to the suboptimal equilibrium either: instead, the system seems to wander around a non-equilibrium state in the proximity of $p = 0.15$. In that state, it is *as if* the disappointment received from instances of coordination failure is too weak to generate behavioral changes. Thus, some behavioral equilibrium (akin to satisficing) emerges at which

players experience too little of an urge to switch to the other strategy, feeling that their chosen behavior works well enough for them. The real benefits from switching to the optimal strategy are simply not large enough at the level of the individual.

**Figure 10.** '*Hi-Lo*' with $p_0$ = 0.1. Series 1: $a$ = 0, Series 2: $a$ = 0.05, $b$ = 1.

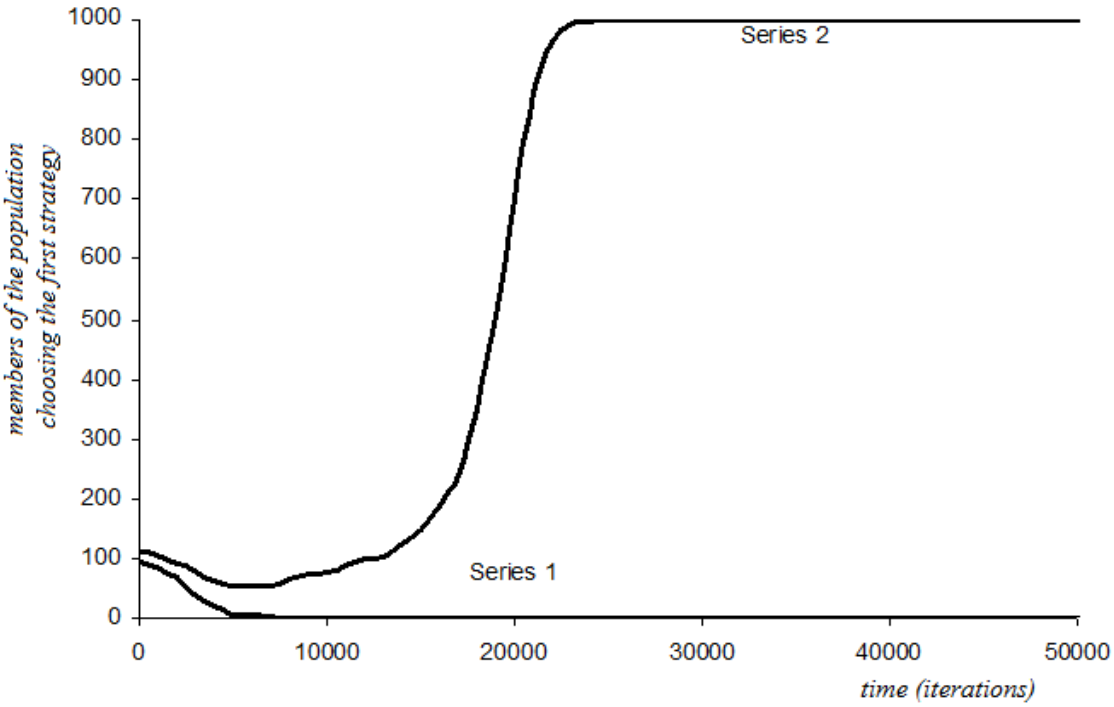



**Figure 11.** '*Stag-Hunt*' with $p_0$ = 0.1. Series 1: $a$ = 0, Series 2: $a$ = 0.1, $b$ = 1.

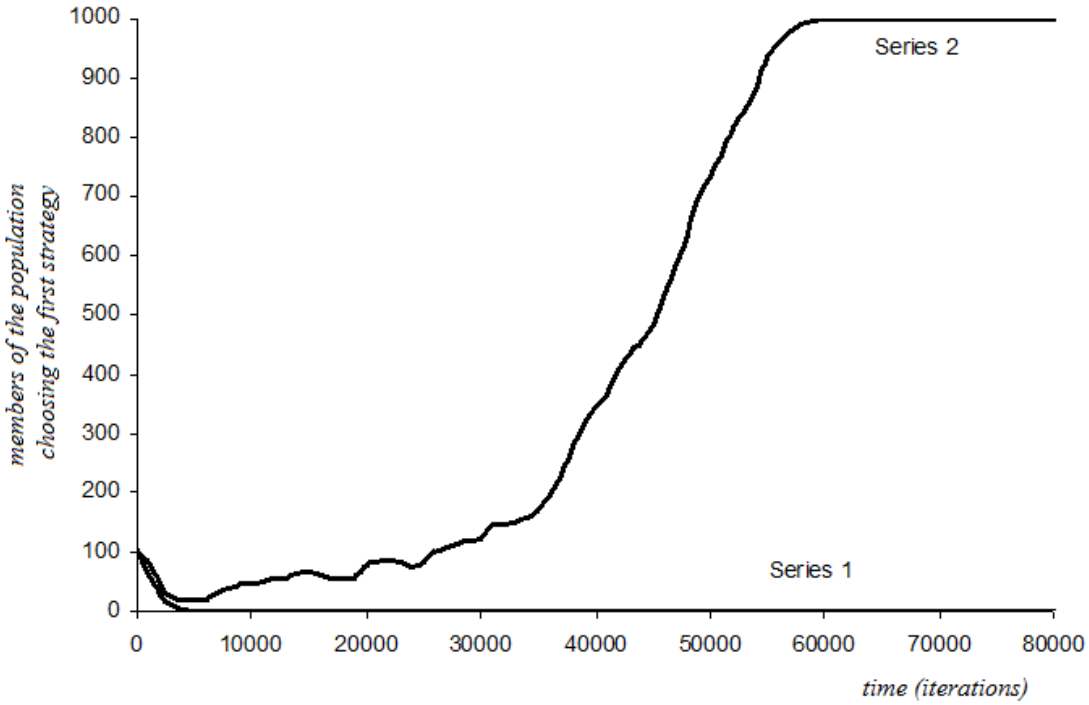
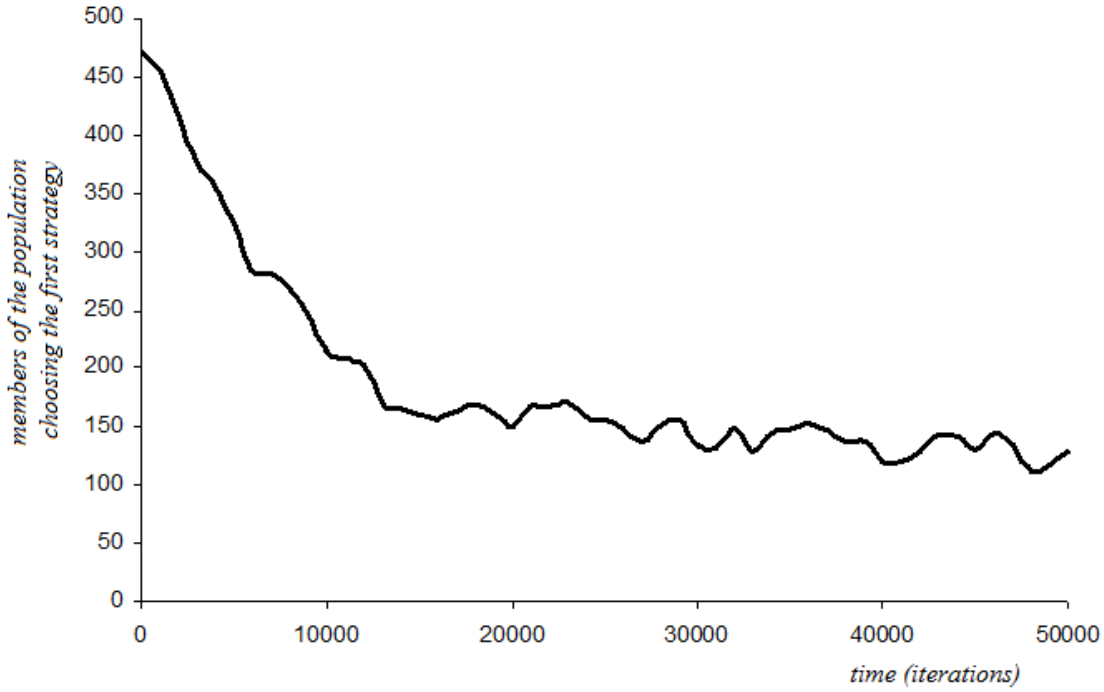


**Figure 12.** '*Stag-Hunt* #3' with $p_0 = 0.5$, $a = 0.25$, $b = 1$.




Tables 4–6 summarize the results of the simulations presented in Subsections 4.6.1 to 4.6.3:

**Table 4.** Simulation results for short-sightedness protocol with stochastic perturbations.

| Protocol: Short-Sightedness | | | |
|---|---|---|---|
| **Game** | **Initial Condition** | **Convergence** | **Comments** |
| Hi-Lo | $p = 0$, $a = 0.01$, $b = 0.1$ | $p = 1$ | A minor perturbation is enough for convergence to $p = 1$. |
| Stag-Hunt | $p = 0.1$, $a = 0.01$, $b = 1$ | $p = 1$ | If $a = 0$ (no perturbations) we have convergence at $p = 0$. |
| Stag-Hunt | $p = 0.8$, $a = 0.01$, $b = 0$ | $p = 0$ | If $a = 0$ (no perturbations) we have convergence at $p = 1$. |
| Stag-Hunt #3 | $p = 0.5$, $a = 0.5$, $b = 1$ | $p = 1$ | If $a = 0$ (no perturbations) we have convergence at $p = 0$. |

**Table 5.** Simulation results for three-period memory protocol with stochastic perturbations.

| Protocol: Three-Period Memory | | | |
|---|---|---|---|
| **Game** | **Initial condition** | **Convergence** | **Comments** |
| Hawk-Dove | $p = 0.5$, $a = 0.33$, $b = 1$ | $p \approx 0.38$ | Convergence at $p \approx 0.47$ in the absence of perturbations. |
| Hi-Lo | $p = 0.35$, $a = 0.1$, $b = 1$ | $p = 0$ | |
| | $p = 0.35$, $a = 0.15$, $b = 1$ | $p = 1$ (6 out of 10 runs) | |
| Stag-Hunt | $p = 0.4$, $a = 0.15$, $b = 1$ | $p = 0$ | Convergence at $p = 1$ under short-sightedness. |
| Stag-Hunt #3 | $p = 0.6$, $a = 0.1$, $b = 0$ | $p = 1$ | If $a = 0$ convergence at $p = 1$ is quicker. |

**Table 6.** Simulation results for additive disappointment protocol with stochastic perturbations.

| Protocol: Additive disappointment | | | |
|---|---|---|---|
| **Game** | **Initial Condition** | **Convergence** | **Comments** |
| Hawk-Dove | $p = 0.5$, $a = 0.1$, $b = 1$ | $p \approx 0.39$ | Convergence at $p \approx 0.41$ in the absence of perturbations. |
| Hi-Lo | $p = 0.1$, $a = 0.05$, $b = 1$ | $p = 1$ | If $a = 0$ (no perturbations) we have convergence at $p = 0$. |
| Stag-Hunt | $p = 0.1$, $a = 0.1$, $b = 1$ | $p = 1$ | If $a = 0$ (no perturbations) we have convergence at $p = 0$. |
| Stag-Hunt #3 | $p = 0.5$, $a = 0.25$, $b = 1$ | $p \approx 0.15$ | |

## 5. Discussion and Conclusions

The simulations presented in the previous section allow for comparisons across the three behavioral 'types' modeled in Section 3 as (1´), (2´), and (3´). None of these 'types' is 'best', in the sense of boosting either individual or social welfare. Each type may perform better in one game and then worse in another. Therefore, in the hypothetical case in which agents have a *choice* as to *their* 'type', it might be optimal (insofar that this is possible) to adopt different 'types' depending on the interaction. For instance, agents may react to disappointment differently in interactions of an antagonistic nature (e.g., '*Hawk-Dove*') from the way they react in cases of coordination failure (e.g., '*Hi-Lo*' or '*Stag-Hunt*'). They may be more rigid in, say, '*Hawk-Dove*' (possibly opting for the additive disappointment protocol) than in pure '*Coordination*', where they may feel more relaxed and conform to the short-sightedness protocol. In fact, the simulation results illustrate that this would indeed be an advantageous tactic, given that none of the examined behavioral codes consistently outperforms the others. Naturally, an interesting extension of this conclusion would be to provide a formal (and quantitative) definition of what constitutes a desired outcome in a game, so that one would be able to calculate the deviation of a specific rule (i.e., behavioral code) from what is thought of as "best".

The paper also illuminated the central role of stochastic perturbations in the determination of the relative social welfare effects of the different behavioral 'types'. Short-sighted agents were shown to be highly sensitive to random perturbations and more likely to benefit from a benevolent third party (a social planner, perhaps?) who directs such shocks in a bid to steering the population to the desired equilibrium. On the other hand, if the planner's intentions are not benign, the population risks being led to a sub-optimal outcome just as easily. In this sense, the three-period memory protocol shields a population from malevolent outside interventions at the expense of reducing the effectiveness of social policy that would, otherwise, have yielded effortless increases in social welfare.

Additive disappointment combines elements from both the short-sightedness and the three-period memory protocols, but its main theoretical disadvantage is that the system seems too sensitive to the choice of two exogenous parameters (the discount factor $\rho$ and the threshold value $Z$). The simulation reported here ($\rho = 0.9$ and $Z = 2\Delta$) implies individuals with a moderate threshold of tolerance, in the sense that it is neither too high to prohibit strategy switches, nor too low to permit too much flexibility. In games featuring multiple evolutionary equilibria, a population of these 'types' may drift somewhere in-between the two equilibria (recall Figure 12). This is consistent with a novel type of behavioral equilibrium which does not correspond to any of the game's evolutionary equilibria. Such a state of behavioral rest is more likely to occur in some form of coordination problem (e.g., '*Stag-Hunt*',

'*Hi-Lo*'), the result being a mixture of behaviors that, while stable, does not correspond to a mixed strategy equilibrium (in the traditional game theoretical sense). To give one real life example, when one observes that QWERTY and DVORAK typewriter keyboards are both still in use, then it is conceivable that this is a behavioral equilibrium state of the type simulated here.

On a similar note, for players acting under additive disappointment, there can be combinations of $\rho$ and $Z$ which yield positive cooperation rates in the '*Prisoner's Dilemma*' or unusually low aggression rates (or even zero) in '*Hawk-Dove*'. This kind of result is consistent with the observation of a stable proportion of 'cooperators' in the '*Prisoner's Dilemma*' (as confirmed by virtually all related experimental work), or players of '*Hawk-Dove*' who bypass opportunities to behave aggressively when their opponents are acting dovishly. The analytical interpretation in this paper is that these players do not find the net benefits from switching to the other strategy high enough to motivate a change in their behavior. While the explanation of why that might be so might lie on bounded rationality considerations, it may also have its roots in the psychological character or social norms pertaining to the agents; e.g., perceptions of fairness or the intrinsic value of cooperation.

Finally, a critical note: This paper has confined its attention to homogeneous populations comprising agents who subscribe exclusively to one of the three revision protocols. A more realistic analysis would allow not only for the coexistence of these protocols in the same population but also for heterogeneity within a single protocol (i.e., agents who, while adopting the additive disappointment protocol, feature different values for parameters $\rho$ and $Z$). Future research along these lines promises to throw important new light on the manner in which learning processes allow populations to achieve greater social and individual success in the games they play.

**Acknowledgements**

I am indebted to Yanis Varoufakis for his valuable comments. I would also like to thank three anonymous reviewers for their helpful suggestions.

**Conflicts of Interest**

The author declares no conflict of interest.

**References**

1. Sugden, R. The evolutionary turn in game theory. *J. Econ. Methodol.* **2001**, *1*, 113–130.
2. Rubinstein, A. *Modeling Bounded Rationality*; MIT Press: Cambridge, MA, USA, 1998.
3. Maynard Smith, J.; Price G.R. The logic of animal conflict. *Nature* **1973**, *246*, 15–18.
4. Maynard Smith, J. The theory of games and the evolution of animal conflicts. *J. Theor. Biol.* **1974**, 47, 209–221.
5. Lewontin, R.C. Evolution and the theory of games. *J. Theor. Biol.* **1961**, *1*, 382–403.
6. Taylor, P.; Jonker L. Evolutionary stable strategies and game dynamics. *Math. Biosci.* **1978**, *40*, 145–156.
7. Mailath, G.J. Do people play Nash equilibrium? Lessons from evolutionary game theory. *J. Econ. Lit.* **1998**, *36*, 1347–1374.

8. Weibull, J.W. What have we learned from evolutionary game theory so far? Research Institute of Industrial Economics IUI, 1998, Working Paper Series 487. Available online: http://swopec.hhs.se/iuiwop/papers/iuiwop0487.pdf (accessed on 1 November 2013).
9. Samuelson, L. Evolution and game theory. *J. Econ. Perspect.* **2002**, *16*, 47–66.
10. Hofbauer, J.; Sigmund K. Evolutionary game dynamics. *B. Am. Math. Soc.* **2003**, *40*, 479–519.
11. Friedman, D. On economic applications of evolutionary game theory. *J. Evol. Econ.* **1998**, *8*, 15–43.
12. Schlag, K.H. Why imitate, and if so, how? A bounded rational approach to multi-armed bandits. *J. Econ. Theory* **1998**, *78*, 130–156.
13. Cross, J. A stochastic learning model of economic behavior. *Q. J. Econ.* **1973**, *87*, 239–266.
14. Borgers, T.; Sarin R. Learning through reinforcement and replicator dynamics. *J. Econ. Theory* **1997**, *77*, 1–14.
15. Erev, I.; Roth, A.E. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* **1998**, *88*, 848–881.
16. Fudenberg, D.; Kreps D. Learning mixed equilibria. *Game Econ. Behav.* **1993**, *5*, 320–367.
17. Fudenberg, D.; Levine D.K. *The Theory of Learning in Games*; MIT Press: Cambridge, MA, USA, 1998.
18. Cressman, R. *Evolutionary Dynamics and Extensive Form Games*; MIT Press: Cambridge, MA, USA, 2003.
19. Sandholm, W.H. *Population Games and Evolutionary Dynamics*; MIT Press: Cambridge, MA, USA, 2010.
20. Foster, D.; Young H.P. Stochastic evolutionary game dynamics. *Theor. Popul. Biol.* **1990**, *38*, 219–232.
21. Kandori, M.; Mailath G.J.; Rob R. Learning, mutation, and long run equilibria in games. *Econometrica* **1993**, *61*, 29–56.
22. Young, H.P. The evolution of conventions. *Econometrica* **1993**, *61*, 57–84.
23. Binmore, K.; Samuelson, L.; Vaughan, R. Musical chairs: Modeling noisy evolution. *Game Econ. Behav.* **1995**, *11*, 1–35.
24. Young, H.P. *Individual Strategy and Social Structure*; Princeton University Press, Princeton, NJ, USA, 1998.
25. Loomes, G.; Sugden, R. Regret theory: An alternative theory of rational choice under uncertainty. *Econ. J.* **1982**, *92*, 805–824.
26. Gul, F. A theory of disappointment aversion. *Econometrica* **1991**, *59*, 667–686.
27. Braun, M.; Muermann, A. The impact of regret on the demand of insurance. *J. Risk Insur.* **2004**, *71*, 737–767.
28. Irons, B.; Hepburn, C.J. Regret theory and the tyranny of choice. *Econ. Rec.* **2007**, *83*, 191–203.
29. Laciana, C.E.; Weber, E.U. Correcting expected utility for comparisons between alternative outcomes. *Journal of Risk and Uncertainty* **2008**, *36*, 1–17.
30. Grant, S.; Atsushi, K.; Polak, B. Different notions of disappointment aversion. *J. Econ. Lit.* **2001**, *81*, 203–208.
31. Hart, S.; Mas-Colell A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* **2000**, *68*, 1127–1150.

32. Macy, M.W.; Flache, A. Learning dynamics in social dilemmas. *P. Natl. A. Sci.* **2002**, *99*, 7229–7236.
33. Hogdson, G.M.; Knudsen T. The complex evolution of a simple traffic convention: the functions and implications of habit. *J. Econ. Behav. Organ.* **2004**, *54*, 19–47.
34. Radax, W.; Wäckerle, M.; Hanappi, H. From agents to large actors and back; Formalized story-telling of emergence and exit in political economy. 2009. Available online: http://publik.tuwien.ac.at/files/PubDat_177962.pdf (accessed on 1 November 2013).
35. Heinrich, T.; Schwardt, H. Institutional inertia and institutional change in an expanding normal-form game. *Games* **2013**, *4*, 398–425.
36. Weibull, J.W. *Evolutionary Game Theory*; MIT Press, Cambridge, MA, USA, 1995.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).