Beatable Imitation in Symmetric Games with Perturbed Payoffs

Tsakas, Nikolas

Universidad Carlos III de Madrid, Singapore University of Technology and Design

12 June 2014

Beatable Imitation in Symmetric Games With Perturbed Payoffs

Nikolas Tsakas

Department of Economics, Universidad Carlos III de Madrid, Spain

Abstract

In a recent paper, Duersch et.al (2012) showed that in a rather broad class of repeated symmetric two-player games, a player who uses the simple "imitate-if-better" heuristic cannot be subject to a money pump. In this paper, we extend the analysis to games with randomly perturbed payoffs and we show that this result is not robust to, even arbitrarily small, payoff perturbations. In particular, we provide a necessary and sufficient condition that characterizes the class of perturbed games in which the imitator can be subject to a money pump.

JEL classification: C72, C73
Keywords: Imitate-if-better, Repeated Games, Symmetric Games, Relative Payoffs, Robustness, Perturbations.

1. Introduction

Imitation is a commonly observed mechanism of decision making in uncertain environments. Some factors that reinforce the use of imitative decision rules are the limited computational capabilities of the players or their limited knowledge about the nature of uncertainty. In the current paper we focus on the learning heuristic described as "imitate-if-better".[1] A player who uses this decision rule adopts the action of his opponent if and only if the opponent's payoff was strictly higher than his in the previous round .

The current paper builds upon the seminal work of Duersch et.al (2012), where the authors show a surprisingly solid behavior of imitation against any sophisticated strategy, in a large class of two-player symmetric games.[2] More specifically, they show that in repeated two-player symmetric games, the simple heuristic "imitate-if-better" can be exploited without bounds if and only if the relative payoff game contains a generalized-rock-paper-scissors submatrix. In any other game, even a truly sophisticated opponent, who may be infinitely patient, forward looking and making no mistakes, cannot receive infinitely higher payoff relatively to the imitator. It is even more surprising that the result stands even

Email address: `ntsakas@eco.uc3m.es` (Nikolas Tsakas)

[1] In our two-player setting, the "imitate-if-better" rule is equivalent with "imitate-the-best".

[2] We refer the reader to this paper for further references related to "imitate-if-better" mechanisms.

when the sophisticated player knows that he faces an imitator and is able to anticipate his future behavior perfectly. Their results cover many games with vast importance in economic applications, such as Cournot duopoly, Bertrand duopoly, public good games and all the symmetric 2×2 games.

However, it turns out that assuming complete information is crucial for the validity of this result. In this paper, we extend the analysis to games with perturbed payoffs and we provide a necessary and sufficient condition that characterizes the class of games where an imitator can be exploited as a money pump by a sophisticated opponent. Indeed, it turns out that the results are not robust to the presence of perturbations to the payoffs. Intuitively, this happens because a sophisticated player can create a probabilistic imitation cycle, where transition between action profiles occurs with strictly positive probability and along the cycle the sophisticated opponent receives strictly higher expected relative payoff compared to the imitator. Under certain conditions, the sophisticated player is able to exploit the imitator even when the perturbations become arbitrarily small.

2. Preliminaries

We present the model and the results obtained in Duersch et.al (2012) which are related to our current analysis.[3]

Consider a symmetric two–player game (X, π), where the agents are endowed with the same (finite) set of pure actions X. Each player has a bounded (stage game) payoff function denoted by $\pi : X \times X \to \mathbb{R}$, where $\pi(x, y)$ denotes the payoff of the player choosing x when her opponent chooses y.

Now, consider the repeated game, where the stage game (X, π) is played in each one of the periods $t = \{0, 1, \ldots, T\}$. The payoff function of the repeated game is equal to the sum of the player's stage game payoffs $\Pi = \sum_{t=0}^{T} \pi(x_t, y_t)$.

Define also the *relative payoff game* (X, Δ) as a symmetric zero–sum game, where, given the symmetric two–player game (X, π), the relative payoff function $\Delta : X \times X \to \mathbb{R}$ is given by:

$$\Delta(x, y) = \pi(x, y) - \pi(y, x)$$

The relative payoff game induces by construction a symmetric zero–sum game, as $\Delta(x, y) = -\Delta(y, x)$.

We call *imitator* the player who follows the simple heuristic "imitate–if–better", meaning that he adopts the action of the opponent if and only if the opponent's payoff was *strictly* higher than his in the previous round.

Definition 1. [*Imitator*] If y_t is the action of the imitator at time t and x_{t-1} the action of the other player at time $t-1$, then

$$y_t = \begin{cases} x_{t-1} & \text{if } \Delta(x_{t-1}, y_{t-1}) > 0 \\ y_{t-1} & \text{if } \textit{otherwise} \end{cases} \tag{1}$$

[3] Slight modifications in notation and terminology are imposed in order to facilitate the introduction the perturbed game.

The other player may use any strategy, may never make mistakes, and may be infinitely patient and forward looking. Moreover, he may know exactly what her opponent will play in every period, including the opponent's starting value. Throughout this paper, we call the second player *sophisticated*.

Definition 2. [*Money pump*] We say that the imitator is subject to a money pump if for some initial action y_0 of the imitator and some sequence $\{x_t\}$ of actions of the opponent it holds that

$$\text{for all } M > 0 \text{ there exists } T > 0 \text{ such that } \sum_{t=0}^{T} \Delta(x_t, y_t) > M$$

where y_t is given by Equation 1.

The above definition is the negation of the definition of "no money pump" given in the original paper. Intuitively, being subject to a money pump means that the opponent can take infinitely large advantage of the imitator in the long run. Otherwise, we say that an imitator is not subject to a *money pump*, i.e. if for all initial actions y_0 of the imitator and all sequences $\{x_t\}$ of actions of the opponent there exists upper bound $M > 0$ such that $\sum_{t=0}^{T} \Delta(x_t, y_t) > M$ for all $T > 0$.

Finally, we define an imitation cycle as follows. In the symmetric two–player game (X, π), a *path* is a sequence of action profiles $(x_0, y_0), (x_1, y_1), \ldots$. A path is called *constant* if for all $t = 0, 1, \ldots$ holds that $(x_t, y_t) = (x_{t+1}, y_{t+1})$; otherwise it is called non-constant. A non-constant finite path $(x_0, y_0), \ldots, (x_l, y_l)$ is a *cycle* if $(x_0, y_0) = (x_l, y_l)$ for some $l > 1$. An *imitation cycle* is a particular cycle where one player always obtains strictly positive relative payoff and the other imitates the action of the first player in the previous round, i.e. for all (x_t, y_t) and (x_{t+1}, y_{t+1}) on the path of the cycle, it holds that $\Delta(x_t, y_t) > 0$ and $y_{t+1} = x_t$.

2.1. Results

We state some of the main results in Duersch et.al (2012). The results are stated without their proofs, for which we refer the reader to the original article.

Lemma 1 (Lemma 1, Duersch et.al (2012))**.** *For any finite symmetric game (X, π), imitation is subject to a money pump if and only if there exists an imitation cycle.*

The intuition behind this lemma is that, unless there exists an imitation cycle, the sophisticated player trying to exploit the imitator will eventually lead the imitator to choose an action that will never give him strictly negative relative payoff, regardless of the choice of her opponent. For example think of the following simple symmetric two–player game.

$$\begin{array}{c c} & \begin{array}{ccc} L & M & H \end{array} \\ \begin{array}{c} L \\ M \\ H \end{array} & \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

In this case, suppose that the lifetime-payoff maximizer (Amy) chooses row and the imitator (Carlos) chooses column. Amy knows that the only action profile that gives her strictly positive relative payoff is (H, L), because $\Delta(H, L) = +1 - (-1) = 2$. Hence, in order to exploit Carlos as a money pump, she must lure him to play L whenever she plays H and this should happen infinitely often. Now, suppose, for example, that the initial action profile is (L, L), so Amy knows that Carlos will stick to action L in the next period. If she tries to exploit him by playing H, then in the following period Carlos will imitate her and play H. From the period that Carlos starts playing H, he never changes back to either L or M, because he never receives lower payoff than Amy. This provides Carlos no temptation to mimic her. The argument is analogous for every initial action profile.

In order to state the main theorem, we need two additional definitions. We say that symmetric zero–sum game (X, π) is a *gerenalized rock–paper–scissors (gRPS) matrix* if for each column there exists a row with strictly positive payoff to the row player, i.e. if for all $y \in X$ there exists an $x \in X$ such that $\pi(x, y) > 0$. Subsequently, a symmetric zero–sum game (X, π) is called a *generalized rock–paper-scissors (gRPS) game* if it contains a submatrix $(\bar{X}, \bar{\pi})$ with $\bar{X} \subseteq X$ and $\bar{\pi}(x, y) = \pi(x, y)$ for all $x, y \in \bar{X}$ and $(\bar{X}, \bar{\pi})$ is a gRPS matrix.

Theorem 1 (Theorem 1,Duersch et.al (2012))**.** *Imitation is subject to a money pump in the finite symmetric game (X, π) if and only if its relative payoff game (X, Δ) is a gRPS game.*

The results reveal surprisingly robust behavior of imitation against other strategies, including very sophisticated ones. The assumption that the imitator sticks to his action in case of a tie in the payoffs is crucial for the validity of the results; a fact that is also mentioned in Duersch et.al (2012). This feature plays an important role during the analysis of the perturbed game.

3. The Perturbed game

In this section, we introduce the symmetric *perturbed game* (X, π^{ε}), where the (stage game) payoff function for any player i, denoted by $\pi^{\varepsilon} : X \times X \to \mathbb{R}$, is given by $\pi^{\varepsilon}(x, y) = \pi(x, y) + \varepsilon(x, y)$, for any given action profile $(x, y) \in X \times X$.[4]

For each $(x, y) \in X \times X$, the *perturbation* $\varepsilon(x, y)$ is a discrete random variable with bounded support and probability distribution $F_{x,y}$.[5] The distributions of the perturbations are common knowledge, but their realized values are observed by the players after they make their choices, so they cannot condition their behavior upon these realizations.

Notice that the game is still symmetric. The perturbation ε depends on the action of the player himself, x, as well as the action of his opponent, y, but it is not different among players. The payoff distribution of row player in the action profile (x, y) is the same as the one of the column player in (y, x).

[4] This is a modification of the "disturbed games" introduced by Harsanyi (1973).

[5] All our results can be extended to continuous random variables.

In the repeated perturbed game the stage game (X, π^{ε}) is played in each one of the periods $t = \{0, 1, \ldots, T\}$. Different realizations of each ε are drawn every period and they are independent and identically distributed across periods.[6] The payoff function of the repeated game is equal to the sum of the player's *realized* stage game payoffs:

$$\Pi^{\varepsilon} = \sum_{t=0}^{T} \pi^{\varepsilon}(x_t, y_t) = \sum_{t=0}^{T} [\pi(x_t, y_t) + \varepsilon(x_t, y_t)]$$

Analogously to Section 2, given the perturbed game (X, π^{ε}), the *perturbed relative payoff game* $(X, \Delta^{\varepsilon})$ is a symmetric zero–sum game with payoff function $\Delta^{\varepsilon} : X \times X \to \mathbb{R}$ given by:

$$\Delta^{\varepsilon}(x, y) = \pi^{\varepsilon}(x, y) - \pi^{\varepsilon}(y, x)$$

Hence, we can define again the behavior of the imitator as in Section 2, noticing that now he imitates based on the *realized* payoffs of the previous period. Namely, for some initial $y_0 \in X$,

$$y_t = \begin{cases} x_{t-1} & \text{if } \Delta^{\varepsilon}(x_{t-1}, y_{t-1}) > 0 \\ y_{t-1} & \text{otherwise} \end{cases} \tag{2}$$

Due to the presence of randomness in the current environment, the definitions need to be slightly modified. Recall that in the baseline game the condition for an imitator to be subject to a money pump was the following:

$$\text{for all } M > 0, \text{ there exists a } T > 0 \text{ s.t. } \sum_{t=0}^{T} \Delta(x_t, y_t) > M.$$

Notice that the previous definition is equivalent to:

$$\text{for all } M > 0, \text{ there exists a } T_M > 0 \text{ s.t. } \sum_{t=0}^{T} \Delta(x_t, y_t) > M \text{ for all } T > T_M.$$

However, in the current probabilistic context the definitions that correspond to each one of the above are not equivalent anymore. We find the second definition more natural, in the sense that the first definition would only make sense if the sophisticated player was able to decide after each period whether to continue the game or not. For this reason we use it to establish our results. It is also easy to see that the second definition implies the first one. Nevertheless, in Section 6, we discuss how the results are modified if we use the alternative definition.

Definition 3. [*Money pump - Perturbed Game*] We say that the imitator is subject to a money pump in the perturbed game, if for some initial action of the imitator y_0 and some sequence $\{x_t\}$ of actions of the opponent:

$$\mathbb{P}[\text{for all } M > 0 \text{ exists } T_M > 0 \text{ s.t. } \sum_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t) > M \text{ for all } T > T_M] = 1$$

where y_t is given by Equation 2.

[6] Not necessarily across action profiles.

This means that almost surely the relative payoffs of the sophisticated player will exceed any given upper bound as long as the game is played for sufficiently many periods.[7]

We then define a *probabilistic imitation cycle* as follows. In the perturbed game a probabilistic path is a sequence of action profiles $(x^0, y^0), (x^1, y^1), ...$ – the superscripts should not be confused with the subscripts that correspond to time–. A probabilistic path is constant if for all $t = 0, 1, \ldots$ and some i it holds that $\mathbb{P}[(x_t, y_t) = (x_{t+1}, y_{t+1}) = (x^i, y^i)] = 1$; otherwise it is called non-constant. A non-constant finite probabilistic path $(x^0, y^0), \ldots, (x^l, y^l)$ is called a probabilistic cycle if $(x^0, y^0) = (x^l, y^l)$ for some $l > 1$, which is called *length* of the cycle.

A *probabilistic imitation cycle* is a particular cycle along which one player makes strictly positive expected relative payoff and the opponent imitates the action of the first player in the previous period with strictly positive probability. Formally,

Definition 4. [*Probabilistic Imitation Cycle*] A *probabilistic imitation cycle is a probabilistic cycle $(x^0, y^0), \ldots, (x^l, y^l)$ that satisfies the following conditions:*

i. *For all (x^i, y^i) and (x^{i+1}, y^{i+1}) on the cycle it holds that $y^{i+1} = x^i$,*
ii. $\mathbb{P}[(x_{t+1}, y_{t+1}) = (x^{i+1}, y^{i+1}) \mid (x_t, y_t) = (x^i, y^i)] > 0,$
iii. $\mathbb{E}\left[\Delta^\varepsilon(x^i, y^i)\right] > 0.$

for all $i = 1, \ldots, l-1$.

Later on, we will show that this definition is sufficient for the imitator being subject to a money pump, but condition (iii) can be relaxed further. Finally, we define the concept of an *ε–gRPS game*

Definition 5. [*ε–gRPS matrix*] A symmetric zero–sum perturbed game (X, π^ε) is called an ε–gRPS matrix if for each column $y \in X$ there exists a row $x \in X$ and a realization of the perturbation $\varepsilon(x, y)$ that occurs with strictly positive probability and yields a strictly positive payoff for the row player, $\pi^\varepsilon(x, y) > 0$.

Subsequently, a symmetric zero–sum game (X, π^ε) is called a *ε–gRPS game* if it contains a submatrix $(\bar{X}, \bar{\pi}^\varepsilon)$ with $\bar{X} \subseteq X$ and $\bar{\pi}^\varepsilon(x, y) = \pi^\varepsilon(x, y)$ for all $x, y \in \bar{X}$ and $(\bar{X}, \bar{\pi}^\varepsilon)$ is a ε–gRPS matrix.

4. Results

Proposition 1. *Consider a finite symmetric perturbed game (X, π^ε) with its relative payoff perturbed game (X, Δ^ε). If there exists a probabilistic imitation cycle, then (X, Δ^ε) is an ε–gRPS game.*

[7] Analogously, we will say that an imitator is not subject to a money pump if for all initial actions of the imitator y_0 and any sequence $\{x_t\}$ of actions of the opponent the probability $\mathbb{P}[\text{exists } M > 0 \text{ s.t. for all } T_M > 0 \text{ exists } T > T_M \text{ s.t. } \sum_{t=0}^{T} \Delta^\varepsilon(x_t, y_t) \leq M] > 0$. In other words, this means that with a strictly positive probability for some $M > 0$ there exists an infinite increasing sequence of time periods in all of which the sum of relative payoffs is bounded by M.

The opposite direction need not be true necessarily. In fact, condition (iii) of Definition 4 is not necessarily satisfied in an ε–gRPS game. Nevertheless, it will be shown that some condition of this type (maybe not that strict) is needed to ensure that an imitator is subject to a money pump. The following proposition establishes a two–direction relation between a ε–gRPS game and a probabilistic imitation cycle.

Proposition 2. *A finite symmetric perturbed relative payoff game (X, Δ^ε) is an ε–gRPS game if and only if it satisfies the conditions (i) and (ii) of Definition 4.*

The following proposition establishes the fact that the existence of an ε–gRPS matrix is a necessary condition for the imitator to be subject to money pump.

Proposition 3. *For imitation to be subject to a money pump the perturbed relative payoff game (X, Δ^ε) needs to be an ε–gRPS game.*

In contrast with Duersch et.al (2012) this is not sufficient anymore.[8] This happens because of the random nature of the setting and will be clarified in the following results.

Theorem 2. *For any finite symmetric perturbed game (X, π^ε), if there exists a probabilistic imitation cycle then imitation is subject to money pump.*

Proof. It is sufficient to show that for some y_0 and a sequence of $\{x_t\}$ it holds that $\mathbb{P}[\sum_{t=0}^{T-1} \Delta^\varepsilon(x_t, y_t) \to +\infty \text{ as } T \to \infty] = 1$ for which it suffices to show that

$$\mathbb{P}\left[\frac{1}{T}\sum_{t=0}^{T-1} \Delta^\varepsilon(x_t, y_t) \to \overline{E\Delta} \text{ as } T \to \infty\right] = 1 \text{ for some } \overline{E\Delta} > 0 \tag{3}$$

To complete the proof we use some standard results from finite Markov chains (see Norris, 1997). First of all, notice that if at any period the players choose one of the action profiles of the probabilistic imitation cycle, then the sophisticated player can ensure that they will stay on the cycle forever. This happen as follows, whenever the imitator receives positive relative payoff the sophisticated player does not change his action, therefore the action profile of the next period remains the same. To the contrary, when the sophisticated player receives positive relative payoff, anticipates the next move of the imitator and plays the action that will lead them to the next action profile on the cycle. Notice also that if the players play always along the cycle, then the process can be represented by a discrete–time, time–homogeneous, finite Markov chain with its state space S being the action profiles of the cycle and the transition probabilities from state $s^i = (x^i, y^i)$ to any other state defined as follows for $i = 1, \ldots, l-1$:[9] with probability $p^{i,i} = \mathbb{P}[\Delta^\varepsilon(x^i, y^i) \leq 0]$ the process remains at the same state and with the remaining probability $p^{i,i+1} = 1 - p^{i,i} = \mathbb{P}[\Delta^\varepsilon(x^i, y^i) > 0]$ the process moves to the state

[8]For the definition of money pump that we use in the main analysis. In Section 6, we show that the game being an ε–gRPS game is also a sufficient condition for what we will call "easy money pump".

[9]Recall that $(x^l, y^l) = (x^1, y^1)$,

$s^{i+1} = (x^{i+1}, y^{i+1})$. It is obvious that this is a recurrent and irreducible Markov chain, with a unique stationary distribution $\mu = (\mu^1, \ldots, \mu^{l-1})$. Among others this means that if the game is played for infinitely many periods then the process will pass infinitely many times from each of the states.

Before we proceed to the proof we need some further definitions – we follow as closely as possible the terminology of Norris (1997) –. For a state s^i, we let $\mathcal{T}_i$ be the first time that the chain visits this state $\mathcal{T}_i := \inf\{t \geq 1 : S_t = s^i\}$, where S^t denotes the state at period t. Analogously, we define as $\mathcal{T}_i^{(r)}$ the r-th passage time to state s^i, $\mathcal{T}_i^{(r)} = \inf\{t > \mathcal{T}_i^{(r-1)} : S_t = s^i\}$, for $r = 2, 3, \ldots$. Subsequently, we define an *excursion* of the process as the trajectory of the chain between two successive visits to the state s^i. Hence, the r–th excursion is defined as follows $\mathcal{X}_i^{(r)} := (S_t, \mathcal{T}_i^{(r)} \leq t < \mathcal{T}_i^{(r+1)})$ for $r = 1, 2, \ldots$ and $\mathcal{X}_0^{(r)} := (S_t, 0 \leq t < \mathcal{T}_i^{(1)})$. Analogously, the length of the r–th excursion to s^i is $\mathcal{S}_i^{(r)} = \mathcal{T}_i^{(r)} - \mathcal{T}_i^{(r-1)}$, if $\mathcal{T}_i^{(r-1)} < \infty$ and equal to 0 otherwise. Finally, we denote the mean return time to state s^i by $m^i = \mathbb{E}_i(\mathcal{S}_i^{(r)})$, which will satisfy $m^i = 1/\mu^i$ (see Norris, 1997, Theorem 1.7.7). Moreover, by the Strong Markov Property (see Norris, 1997, Lemma 1.4.2) both the excursions starting from a state and their lengths (as functions of them) are independent and identically distributed random variables. Therefore, if we denote by $V^i(T)$ the number of times that the process passes from state s^i during the first T periods and we use the strong law of large numbers then (for a detailed proof see Norris, 1997, Theorem 1.10.2)

$$\mathbb{P}\left[\frac{V^i(T-1)}{T} \to \mu^i \text{ as } T \to \infty \text{ for all } i\right] = 1 \tag{4}$$

Furthermore, conditional on being at state (x^i, y^i) the relative payoffs $\Delta^\varepsilon(x^i, y^i)$ are i.i.d. random variables. Moreover, recall that all the states are recurrent, which means $\mathbb{P}[V^i(T) \to \infty \text{ as } T \to \infty] = 1$. This together the strong law of large numbers imply that

$$\mathbb{P}\left[\frac{1}{V^i(T)} \sum_{t=\mathcal{T}_i^{(1)}}^{\mathcal{T}_i^{(V^i(T))}} \Delta^\varepsilon(x_t, y_t) \to \mathbb{E}\left[\Delta^\varepsilon(x^i, y^i)\right] \text{ as } T \to \infty \text{ for all } i\right] = 1 \tag{5}$$

Consider now the quantity $\overline{E\Delta} = \sum_{i=1}^{l-1} \mu^i \mathbb{E}[\Delta^\varepsilon(x^i, y^i)]$, where $\overline{E\Delta} > 0$ for sure because $\mathbb{E}[\Delta^\varepsilon(x^i, y^i)] > 0$ for all i by condition (iii) of Definition 4 (One can easily see why this condition can be relaxed).

$$\left|\frac{1}{T}\sum_{t=0}^{T-1} \Delta^\varepsilon(x_t, y_t) - \overline{E\Delta}\right| = \left|\sum_{i=1}^{l-1}\left(\frac{V^i(T-1)}{T}\frac{1}{V^i(T-1)}\sum_{t=\mathcal{T}_i^{(1)}}^{\mathcal{T}_i^{(V^i(T-1))}} \Delta^\varepsilon(x_t, y_t) - \mu^i \mathbb{E}[\Delta^\varepsilon(x^i, y^i)]\right)\right| \leq$$

$$\leq \sum_{i=1}^{l-1} \left(\left| \frac{V^i(T-1)}{T} - \mu^i \right| \left| \frac{1}{V^i(T-1)} \sum_{t=\mathcal{T}_i^{(1)}}^{\mathcal{T}_i^{(V^i(T-1))}} \Delta^\varepsilon(x_t, y_t) - \mathbb{E}[\Delta^\varepsilon(x^i, y^i)] \right| \right) +$$

$$+ \sum_{i=1}^{l-1} \left(\mu^i \left| \frac{1}{V^i(T-1)} \sum_{t=\mathcal{T}_i^{(1)}}^{\mathcal{T}_i^{(V^i(T-1))}} \Delta^\varepsilon(x_t, y_t) - \mathbb{E}[\Delta^\varepsilon(x^i, y^i)] \right| + \mathbb{E}[\Delta^\varepsilon(x^i, y^i)] \left| \frac{V^i(T-1)}{T} - \mu^i \right| \right)$$

Given $\delta > 0$, from Equations 4 and 5 there exists $T_i(\delta)$ for each i such that for all $T > T_i(\delta)$: $\left| \frac{1}{V^i(T-1)} \sum_{t=\mathcal{T}_i^{(1)}}^{\mathcal{T}_i^{(V^i(T-1))}} \Delta^\varepsilon(x_t, y_t) - \mathbb{E}[\Delta^\varepsilon(x^i, y^i)] \right| < \max\left\{ \sqrt{\frac{\delta}{2l}}, \frac{\delta}{4l} \right\}$ and $\left| \frac{V^i(T-1)}{T} - \mu^i \right| < \max\left\{ \frac{\delta}{4\overline{E}l}, \sqrt{\frac{\delta}{2l}} \right\}$, where $\overline{E} = \max_{i=1,\dots,l} \mathbb{E}[\Delta^\varepsilon(x^i, y^i)]$.

Therefore, denoting $\overline{T} = \max_{i=1,\dots,l} T_i(\delta)$ and given that $\mu^i \leq 1$ then for all $T > \overline{T}$:

$$\left| \frac{1}{T} \sum_{t=0}^{T-1} \Delta^\varepsilon(x_t, y_t) - \overline{E\Delta} \right| \leq \sum_{i=1}^{l-1} \left(\sqrt{\frac{\delta}{2l}} \sqrt{\frac{\delta}{2l}} + \mu^i \frac{\delta}{4l} + \mathbb{E}[\Delta^\varepsilon(x^i, y^i)] \frac{\delta}{4\overline{E}l} \right) \leq$$

$$\leq \sum_{i=1}^{l-1} \left(\frac{\delta}{2l} + \frac{\delta}{4l} + \frac{\delta}{4l} \right) < \delta$$

Therefore, we have shown that $\mathbb{P}\left[\frac{1}{T} \sum_{t=0}^{T-1} \Delta^\varepsilon(x_t, y_t) \to \overline{E\Delta} \text{ as } T \to \infty \right] = 1$. $\square$

The result provides a sufficient condition for the imitator to be subject to a money pump. Nevertheless, this condition might not be necessary. For this reason, our next step is to establish a necessary and sufficient condition for imitation to be subject to a money pump. The following definition is a weaker version of the probabilistic imitation cycle, where the main difference is that it allows the cycle to contain states where the sophisticated player does not receive positive expected relative payoffs. These states, although harmful for the sophisticated player, might be necessary for allowing him to complete the cycle and if the proportion of time that the players spend on them is sufficiently low, then the imitator can still be exploited.

Definition 6. [*Weak Probabilistic Imitation Cycle*] A *weak probabilistic imitation cycle is a probabilistic cycle* $(x^0, y^0), \dots, (x^l, y^l)$ *that satisfies the following conditions:*

i. *For all* (x^i, y^i) *and* (x^{i+1}, y^{i+1}) *on the cycle it holds that* $y^{i+1} = x^i$,

ii. $\mathbb{P}[(x_{t+1}, y_{t+1}) = (x^{i+1}, y^{i+1}) \mid (x_t, y_t) = (x^i, y^i)] > 0$, *for all* $i = 1, \dots, l-1$ *and*

iii. $\sum_{i=1}^{l-1} \mu^i \mathbb{E}\left[\Delta^\varepsilon(x^i, y^i) \right] > 0$.

where μ^i *is the proportion of periods action profile* (x^i, y^i) *is played in the long–run.*

The next corollary follows immediately from the fact that conditions (i) and (ii) of Definitions 6 and 4 are the same and the result established by Proposition 2.

Corollary 1. *Consider a finite symmetric perturbed game (X, π^{ε}) with its relative payoff perturbed game $(X, \Delta^{\varepsilon})$. If there exists a weak probabilistic imitation cycle, then $(X, \Delta^{\varepsilon})$ is an ε–gRPS game.*

Theorem 3. *For any finite symmetric perturbed game (X, π^{ε}), imitation is subject to money pump if and only if there exists a weak probabilistic imitation cycle.*

Proof. ($\Rightarrow$) The proof of this direction is identical to that of Theorem 2 if one observes that the quantity $\overline{E\Delta}$ is still strictly positive if we use condition (iii) of Definition 6 instead of that of Definition 4.

($\Leftarrow$) This direction is proven by contradiction. Consider a finite symmetric perturbed game in which imitation is subject to a money pump and suppose that there exists no weak probabilistic imitation cycle in this game. First, notice that condition (i) of Definition 6 can hold by construction, so non–existence of weak probabilistic imitation cycle might occur for two reasons.

On the one hand, if condition (ii) of Definition 6 fails for all possible cycles that satisfy condition (i). In this case, the game cannot be an ε–gRPS game, hence the imitator cannot be subject to a money pump (as it was proven in Proposition 3), which contradicts the premise.

On the other hand, if for any probabilistic cycle that condition (ii) of Definition 6 holds, condition (iii) of the same definition fails. Namely, if there is strictly positive probability to move from one action profile of the cycle to the next one, but expected profits along the cycle are not positive, i.e. $\sum_{i=1}^{l-1} \mu^i \mathbb{E}\left[\Delta^{\varepsilon}(x^i, y^i)\right] \leq 0$. In this case, if the inequality is strict, then in fact this creates a weak probabilistic imitation cycle in favor of the imitator and if the sophisticated player decides to stay on the cycle he will be subject to a money pump himself, which contradicts the initial argument. Similarly, if $\sum_{i=1}^{l-1} \mu^i \mathbb{E}\left[\Delta^{\varepsilon}(x^i, y^i)\right] = 0$, then notice that the very same conditions hold for both players. Therefore, if along the cycle the imitator is subject to a money pump, then the same should hold for the sophisticated player. But this cannot be possible because the game is a zero–sum game. □

Concluding, notice that our results do not require a minimum size of perturbations. Therefore, as long as the conditions for the existence of (weak) probabilistic imitation cycle are satisfied then the results hold even if the size of perturbations is arbitrarily small. This means that we can construct sequences of perturbed games, with decreasing size of perturbations, that converge to an unperturbed game and are such that in the whole sequence of games the imitator is subject to a money pump, whereas in the limiting unperturbed game he is not.

5. An Example

Consider the symmetric two–player game in the example of Section 2, played by Carlos (an imitator – column player) and Amy (a sophisticated player – row player), who knows that she faces an imitator. The set of actions is $X = \{L, M, H\}$ is the same as before and the payoffs are given by the following matrix:

$$\begin{array}{c} \\ L \\ M \\ H \end{array} \begin{array}{c} \begin{array}{ccc} L & M & H \end{array} \\ \begin{pmatrix} 0 & -2\varepsilon & -1 \\ 2\varepsilon & 0 & \varepsilon \\ 1 & -\varepsilon & 0 \end{pmatrix} \end{array}$$

where the perturbation $\varepsilon : \Omega \to \mathbb{R}$ is a random variable. The sample space of ε is $\Omega = \{s, f\}$, with typical element ω, where s stands for success and occurs with probability $p_s = 1/3$ and f stands for failure and occurs with the remaining probability $p_f = 2/3$. The value of the perturbations is equal to

$$\varepsilon = \begin{cases} \delta & \text{if } \omega = s \\ -\delta & \text{if } \omega = f \end{cases}$$

where $\delta \geq 0$.

For $\delta = 0$ there are no perturbations and the game is identical to the one in Section 2, where one can easily observe that this game is not a $gRPS$–game. Therefore, by Theorem 1 (shown by Duersch et.al, 2012) in the unperturbed game Carlos (the imitator) cannot be subject to a money pump. In fact, for any initial action profile, Amy (the sophisticated) can gain a relative payoff at most equal to her maximum one–shot relative payoff, i.e. $\Delta(H, L) = 1 - (-1) = 2$.

For any $\delta > 0$ the payoffs are randomly perturbed according to the realization of the random variable ε. The first thing one should notice is that the game is still symmetric. Subsequently, we check whether there exists a (weak) probabilistic imitation cycle. Consider the following cycle:

$$(H, L) \to (M, H) \to (L, M) \to (H, L)$$

First of all, it is apparent that along the cycle Carlos imitates the action of Amy, which is the first condition that characterizes both a probabilistic and a weak probabilistic imitation cycle. Furthermore, it is also easy to see that in each of the three action profiles Amy receives strictly positive relative payoff with strictly positive probability. This in turn means that the players move from each action profile of the cycle to the next one with a strictly positive probability, which is the second condition for both a probabilistic and a weak probabilistic imitation cycle. In particular, the exact probabilities of passing from each action profile to the next one are shown in the following figure.

$$\begin{aligned} (H, L) \xrightarrow{1} (M, H) \xrightarrow{2/3} (M, H) & \\ \xrightarrow{1/3} (L, M) \xrightarrow{1/3} (L, M) & \\ \xrightarrow{2/3} (H, L) & \end{aligned}$$

Nevertheless, if we calculate the expected relative payoffs for the three profiles we find that $\mathbb{E}[\Delta^\varepsilon(H, L)] = 2$, $\mathbb{E}[\Delta^\varepsilon(M, H)] = -\delta/3$ and $\mathbb{E}[\Delta^\varepsilon(L, M)] = 2\delta/3$.

Hence, given that $\mathbb{E}[\Delta^{\varepsilon}(M,H)] = -\delta/3 < 0$ this is not a probabilistic imitation cycle. Therefore, we cannot yet conclude whether the imitator could be subject to money pump. But is it a weak probabilistic imitation cycle?

In order to answer this we need to construct the relevant Markov chain and calculate its stationary distribution. Doing so we obtain the following transition matrix:

$$\begin{array}{c} \\ (H,L) \\ (M,H) \\ (L,M) \end{array} \begin{array}{c} \begin{array}{ccc} (H,L) & (M,H) & (L,M) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ 2/3 & 0 & 1/3 \end{pmatrix} \end{array}$$

which has stationary distribution $\mu = (\mu^1, \mu^2, \mu^3) = (2/11, 6/11, 3/11)$. Therefore, the quantity $\sum_{i=1}^{l-1} \mu^i \mathbb{E}\left[\Delta^{\varepsilon}(x^i, y^i)\right] = 2\frac{2}{11} - \frac{\delta}{3}\frac{6}{11} + \frac{2\delta}{3}\frac{3}{11} = 4/11 > 0$, for any $\delta > 0$. This verifies that the given sequence of action profiles creates a weak probabilistic imitation cycle. Therefore, in this example Carlos is subject to a money pump.

Finally, notice that the result holds for any $\delta > 0$, even if it becomes arbitrarily small, but it does not hold anymore if $\delta = 0$. Hence, one can conclude that in a certain class of games "unbeatability" of imitation is not robust to arbitrarily small perturbations in the payoffs.

6. Discussion

6.1. *Alternative Definition of Money Pump*

As we have already mentioned there is an alternative definition of money pump that makes exploitation of the imitator easier in perturbed games. In the baseline game the deterministic analog to this definition was equivalent to the one we use in our analysis. This definition is as follows:

Definition 7. [*Easy Money pump - Perturbed Game*] We say that the imitator is subject to a money pump in the perturbed game, if for some initial action of the imitator y_0 and some sequence $\{x_t\}$ of actions of the opponent:

$$\mathbb{P}[\text{for all } M > 0 \text{ exists } T > 0 \text{ s.t. } \sum_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t) > M] = 1$$

where y_t is given by Equation 2.

This means that an imitator is not subject to an easy money pump if for all initial actions of the imitator y_0 and any sequence $\{x_t\}$ of actions of the opponent $\mathbb{P}[\text{exists } M > 0 \text{ s.t. } \sum_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t) \leq M \text{ for all } T \geq 0] > 0$. Notice that this definition makes sense only when the sophisticated player can choose after each period whether to continue playing the game or not.

Moreover, as one can observe this definition favors the sophisticated player, since it is more easily satisfied. In fact, under this revised definition the existence of an ε–gRPS matrix is not only necessary but also sufficient for the imitator to be subject to an easy money pump.

Theorem 4. *For any finite symmetric perturbed game (X, π^{ε}), imitation is subject to easy money pump if and only if the game is an ε–gRPS game.*

The idea behind this result is that if the game lasts long enough, then eventually for any upper bound M it will be realized a sufficiently long sequence of positive realizations for the sophisticated player that will make his profits strictly larger than M.

7. Conclusion

We have shown that imitation of successful behavior in repeated games is vulnerable to exploitation in case that the payoffs of the game are perturbed. Under certain conditions, the result holds even when the perturbations are arbitrarily small and the perturbed game converges to the game of complete information.

This result points out the fact that imitation being "unbeatable", in the sense of not being subject to a money pump, is not robust to the presence of uncertainty regarding the payoffs of the game. Small fluctuations can allow a sophisticated player to exploit the naivety of the imitator.

Appendix

Proof of Proposition 1. The proof is by contradiction. Suppose that there exists a probabilistic imitation cycle with action profiles $(x^0, y^0), \ldots, (x^l, y^l)$. We will show that the sub–matrix $(\bar{X}, \Delta^{\varepsilon})$ is an ε–gRPS matrix, where $\bar{X} = \{x^0, \ldots, x^{l-1}\}$. First of all, notice that condition (i) of Definition 4 implies that along the cycle the set of actions for the row and the column players coincide, i.e. $\{x^0, \ldots, x^{l-1}\} = \{y^0, \ldots, y^{l-1}\}$. Consider the sub–matrix defined by the set of actions $\bar{X}$ and the respective relative payoffs and assume that there exists a column action $y \in \bar{X}$ such that for all row actions $x \in \bar{X}$ and all realizations $\varepsilon(x, y)$ it holds that $\Delta^{\varepsilon}(x, y) \leq 0$. Without loss of generality assume $y = y^0$. By definition of the imitator's behavior (Equation 2) it must hold that $y_{t+1} = y_t$ for all $t \geq 0$. On the one hand, if $y_t \neq x_t$, we have arrived to a contradiction of condition (ii) because the probability of observing the next action profile on the cycle is equal to zero. On the other hand, if $y_t = x_t$ for all $t \geq 0$, then the probabilistic path is constant, which contradicts the definition of probabilistic imitation cycle as being a particular non–constant probabilistic path. $\square$

Proof of Proposition 2. We will prove that if $(X, \Delta^{\varepsilon})$ is an ε–gRPS game then it satisfies conditions (i) and (ii) of Definition 4. The converse follows directly from the fact that condition (iii) was not used at all in the proof of Proposition 1. Consider a path starting from the action profile (x^0, y^0) that belongs to the ε–gRPS sub–matrix of the ε–gRPS game, where x^0 is such that there exists realization of $\varepsilon(x^0, y^0)$ for which $\Delta^{\varepsilon}(x^0, y^0) > 0$. Such action should exist by definition of ε–gRPS game. Now let (x^1, y^1) be such that $y^1 = x^0$ and x^1 be such that there exists realization of $\varepsilon(x^1, y^1)$ that occurs with strictly positive probability for which $\Delta^{\varepsilon}(x^1, y^1) > 0$. Once again such action necessarily exists. Notice that the action profile (x^1, y^1) satisfies conditions (i) and (ii) of Definition 4; condition (i) because $y^1 = x^0$ and condition (ii) because $\Delta^{\varepsilon}(x^0, y^0) > 0$ will occur with strictly positive probability leading the imitator to choose $y^1 = x^0$ and the sophisticated player x^1. Analogously, we construct $(x^2, y^2), (x^3, y^3), \ldots$ that all satisfy the two conditions. The fact that the game is an ε–gRPS game ensures that such action profiles exist. Given that the set of actions is finite, the constructed probabilistic path cannot consist of infinitely many distinct action profiles, therefore there must exist a probabilistic cycle $(x^0, y^0), \ldots, (x^l, y^l)$, where $(x^0, y^0) = (x^l, y^l)$. Moreover, all action profiles of this cycle satisfy conditions (i) and (ii) of Definition 4, hence it is a probabilistic imitation cycle. $\square$

Proof of Proposition 3. The proof is by contradiction. Suppose that $(X, \Delta^{\varepsilon})$ is not an ε–gRPS game. This means that for any sub–matrix $(\bar{X}, \Delta^{\varepsilon})$ of the game, there exists at least one column $y \in \bar{X}$ such that $\Delta^{\varepsilon}(x, y) \leq 0$ for all rows $x \in \bar{X}$ and all realizations of $\varepsilon(x, y)$. We will show by induction that there is no sequence of action profiles that can lead the imitator to be subject to a money pump. Suppose that there exists such sequence and that it includes all the available actions of the action set, i.e. $\{x^1, x^2, \ldots x^n\} = X$. Without loss of generality assume that action x^1 satisfies the above described condition. If at any period T it happens that $y_T = x^1$, then $y_{T'} = x^1$ for all $T' > T$, leading the other player to non–positive profits from period T onwards. Therefore, if this happens the sum of profits for the sophisticated player would be bounded above by $\max\limits_{\varepsilon(x_t, y_t), \forall t} \{\sum\limits_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t)\}$ which means that the imitator cannot be subject to a money pump. Hence, for any $T > 0$ the probability that the sophisticated player plays action x^1 in any of the periods should be equal to zero, because this will lead the imitator to play it at the following period. Therefore, if there exists a sequence of action profiles for which the imitator is subject to money pump, this should not contain x^1, so it should contain at most all actions in $X \setminus \{x^1\} = \{x^2, \ldots x^n\}$. To complete the induction argument, suppose that there exists such sequence and contains actions in $\{x^i, \ldots, x^n\}$, for some $1 < i \leq n$. Then, given that the game is not an ε–gRPS game we can use the same argument as before. Then there must exist a column action, w.l.o.g assume it is x^i, that should never be played, hence the sequence can contain at most actions $\{x^{i+1}, \ldots, x^n\}$. The fact that the set of actions is finite, leads to the conclusion that the set of all actions that can be played in a sequence such that the imitator is subject to money pump is the empty set. Equivalent saying, when the perturbed relative payoff game $(X, \Delta^{\varepsilon})$ is not an ε–gRPS game there is no sequence of actions of the sophisticated player for which the imitator is subject to a money pump. $\square$

Proof of Theorem 4. $(\Rightarrow)$ The proof is identical to that of Proposition 3.
$(\Leftarrow)$ The key point for this part is that it is enough to exist one period for each upper bound at which the sophisticated player has accumulated higher payoff than this bound.

The existence of an ε–gRPS matrix means that along that cycle in each period the sophisticated player receives positive relative payoff with strictly positive probability. Therefore, for each $M > 0$ there always exists a sufficiently long sequence of realized perturbations that favor the sophisticated player. Hence, there always exists $T_M = \min\{T : \mathbb{P}[\sum_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t) > M] > \delta\}$. This also implies the existence of $T'_M = \min\{T : \mathbb{P}[\sum_{t=T_M+1}^{T'_M} \Delta^{\varepsilon}(x_t, y_t) > M - A \mid \sum_{t=0}^{T_M} \Delta^{\varepsilon}(x_t, y_t) = A \leq M] > \delta\}$ for any A, where A depends on the realized perturbation. In fact, each realized history induces sequences $A^{\infty}_M = \{A^1_M, A^2_M, \dots\}$ and $T^{\infty}_M = \{T^1_M, T^2_M, \dots\}$ that satisfy the above conditions. We denote by $\mathcal{A}_M$ and $\mathcal{T}_M$ the sets of all possible sequences that can arise for different histories of realized perturbations. Moreover, the events $\{\sum_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t) > M\}$ and $\{\sum_{t=T_M+1}^{T'_M} \Delta^{\varepsilon}(x_t, y_t) > M - A \mid \sum_{t=0}^{T_M} \Delta^{\varepsilon}(x_t, y_t) = A \leq M\}$ are conditionally independent.

Suppose that the imitator is not subject to an easy money pump and let $M > 0$ be the upper bound that satisfies the no money pump condition, then

$$
\begin{aligned}
&\mathbb{P}\left[\sum_{t=0}^{T} \Delta^{\varepsilon}(x_t, y_t) \leq M \text{ for all } T \geq 0\right] = \\
=&\mathbb{P}\left[\bigcup_{\substack{A^{\infty}_M \in \mathcal{A} \\ T^{\infty}_M \in \mathcal{T}_M}} \left[\bigcap_{k=2}^{\infty} \left(\sum_{T^k_M+1}^{T^{k+1}_M} \Delta^{\varepsilon}(x_t, y_t) > M - A^k_M \mid \sum_{0}^{T^k_M} \Delta^{\varepsilon}(x_t, y_t) = A^k_M \leq M\right) \cap \left(\sum_{t=0}^{T^1_M} \Delta^{\varepsilon}(x_t, y_t) \leq M\right)\right]\right] \\
=& \sum_{\substack{A^{\infty}_M \in \mathcal{A} \\ T^{\infty}_M \in \mathcal{T}_M}} \left[\mathbb{P}\left[\bigcap_{k=2}^{\infty} \left(\sum_{T^k_M+1}^{T^{k+1}_M} \Delta^{\varepsilon}(x_t, y_t) > M - A^k_M \mid \sum_{0}^{T^k_M} \Delta^{\varepsilon}(x_t, y_t) = A^k_M \leq M\right) \cap \left(\sum_{t=0}^{T^1_M} \Delta^{\varepsilon}(x_t, y_t) \leq M\right)\right]\right] \\
=& \sum_{\substack{A^{\infty}_M \in \mathcal{A} \\ T^{\infty}_M \in \mathcal{T}_M}} \left[\mathbb{P}\left[\sum_{t=0}^{T^1_M} \Delta^{\varepsilon}(x_t, y_t) \leq M\right] * \prod_{k=2}^{\infty} \mathbb{P}\left[\sum_{T^k_M+1}^{T^{k+1}_M} \Delta^{\varepsilon}(x_t, y_t) > M - A^k_M \mid \sum_{0}^{T^k_M} \Delta^{\varepsilon}(x_t, y_t) = A^k_M \leq M\right]\right] \leq \\
\leq& \sum_{\substack{A^{\infty}_M \in \mathcal{A} \\ T^{\infty}_M \in \mathcal{T}_M}} \left[\lim_{k \to \infty} (1 - \delta)^k\right] = 0
\end{aligned}
$$

which completes the argument, since for every M the probability of never reaching a period for which the accumulated relative payoff of the sophisticated player is larger than M goes to zero as time goes to infinity. □

Acknowledgments

I am indebted to Antonio Cabrales, Marco Celentani, Daniel Garcia, Jesper Rudiger, Christos Mavridis, Peter Eccles and Marco Serena for fruitful conversations and useful comments that improved significantly the quality of this paper.

I gratefully acknowledge the financial help of the Spanish Ministry of Science and Technology under grant BES-2012-054732. All remaining errors are my sole responsibility.

References

DUERSCH, P., OECHSSLER, J. AND SCHIPPER, B.C. (2012). Unbeatable Imitation, *Games Econ Behav* 76(1), 88–96.

FUDENBERG, D. & TIROLE, J. (1991). *Game Theory*. The MIT Press, Cambridge, MA.

HARSANYI, J. (1973). Games with randomly disturbed payoffs: A new rationale for mixed strategy equilibrium points. *Int J Game Theory* 2(1), 1–23.

NORRIS J.R. (1997). *Markov Chains*. Cambridge University Press, New York, NY.