Agility in Repeated Games: An Example*

Ran Spiegler†

March 3, 2015

Abstract

I study a repeated matching-pennies game between players having limited "agility": when player i decides to switch his action, it takes (geometrically distributed) time for the decision to be implemented. I characterize the unique Nash equilibrium in this game when the papers are sufficiently agile. Players obtain the same equilibrium payoff as in the benchmark game with unlimited agility. However, equilibrium behavior displays endogenous hysteresis, which is more pronounced for less agile players.

Keywords: guerilla, repeated games, imperfect monitoring, agility, organizational behavior, hysteresis

1 Introduction

Conventional economic theory models long-run interaction between individual players and organizations using essentially the same tools. Yet there are important differences in the decision processes that characterize individuals and organizations in such settings. For example, when a military organization chooses to redeploy its forces or change its fighting strategy, implementing this decision takes time. Similarly, when the leadership of a large business organization reaches a strategic product-design decision, it takes time for the decision to trickle down the organization's hierarchy.

Two important real-life examples come to mind. Imagine a large regular army fighting against a guerilla military organization. When the two armies fight in the same arena, the large army is likely to win thanks to its superior means. However, the same asymmetry also implies that if the guerilla organization switched to another arena,

*This project has been supported by The Sapir Center for Development. I thank Johannes Horner for a helpful discussion.

†Tel Aviv University, University College London and CFM. E-mail: rani@post.tau.ac.il.

the regular army could only slowly adapt to the change. The guerilla army's smallness is a disadvantage in head-to-head conflicts, but it is an advantage in terms of agility. Similarly, imagine a software giant competing against a small start-up for a certain market segment. When market demand is stable, the giant's superior infrastructure and marketing capabilities will enable it to smash the start-up. However, if features of market demand change, the start-up may be able to adapt its product more quickly.

This note is a small step toward a model of dynamic games between diversely agile organizations, where agility is measured by the amount of time it takes the organization to implement a change in its strategy. I construct a model of repeated interaction between a strong, heavy organization and a weak, agile organization. As an example, I focus on *Repeated Matching Pennies*. The modeling innovation is that I introduce a new parameter for each player, referred to as his agility. Whenever a player chooses to switch his action, implementing this change takes time, according to a geometric distribution. Player i's agility is thus captured by the single parameter that characterizes this geometric distribution, namely the per-period implementation probability β_i. The more agile player is assumed to be the one who benefits from miscoordination. This captures the idea that strength and heaviness are positively correlated, such that if two organizations fight in the same "arena", the heavier organization will prevail thanks to its superior strength. Finally, I assume that each player can perfectly monitor his opponent's past actions, but he does not observe the opponent's decision to change his course of action. Thus, the model falls into the general category of games with imperfect private monitoring.

I characterize Nash equilibrium in this game, under the assumption that $\beta_1, \beta_2 \geq \frac{1}{2}$ - i.e., both players are sufficiently agile. I show that in this case each player has a strategy that induces the stationary observed switching rate of $\frac{1}{2}$. Therefore, the value of the game is exactly as in the benchmark model, such that superior agility does not lead to a payoff advantage; and the equilibrium behavior induces a stationary observed switching probability of $\frac{1}{2}$. However, the equilibrium probability of player i's *decision* to switch is not stationary, but strictly *decreasing* in the number of periods that elapsed since the last realized switch, at a constant rate of β_i per period. Thus, the players exhibit *hysteresis* in equilibrium - their tendency to stick to their current action becomes stronger over time, and in the limit it becomes overwhelming. Moreover, *the less agile player exhibits greater hysteresis.* This effect is consistent with the intuition that larger (i.e. heavier) organizations are also more "conservative", in the sense that they are less likely to change their behavior over time.

There are a few strands of related literature. Lipman and Wang (2000) study repeated games with switching costs, which is an alternative way to capture limited agility. The present model is also related to the literature on repeated games with observation lags - see Scarf and Shapley (1957), Bhaskar and Obara (2011), Fudenberg et al. (2014), and the references therein. Finally, the general idea of modeling repeated games between *organizations* is prefigured by the literature on endogenous complexity of repeated-game strategies (Rubinstein (1986), Abreu and Rubinstein (1988)).

2 A Model

Two players, 1 and 2, play an infinitely repeated zero-sum game with a common discount factor δ. The set of actions for each player is $\{A, B\}$. Let a_i^t denote player i's chosen action at period $t = 1, 2, ...$ Player 1's stage-game payoffs are as follows: $u_1(a_1, a_2) = 1$ (0) when $a_1 = a_2$ ($a_1 \neq a_2$).

The departure from the conventional repeated-game model is the following. Suppose that at period t, player i chooses $a_i^t \neq a_i^{t-1}$. Then, this decision to change course is not implemented instantaneously. Rather, for every $k = 0, 1, 2, ...$, the switch is implemented at period $t + k$ with probability $\beta_i(1 - \beta_i)^k$, where $\beta_i \geq \frac{1}{2}$. I assume that the player is unable to take any action from the moment a decision to switch has been made until the moment the decision has been implemented. If the switch is implemented at some period $t' > t$, the player is free to choose to switch again at $t' + 1$. Fix an arbitrary initial condition (a_1^0, a_2^0).

I assume that each player perfectly monitors the opponent's past actions and his own entire private history. Thus, when player i observes that player j has played the same action in a number of consecutive periods, he does not know whether this reflects the slow implementation of a past decision to switch, or the fact that a decision to switch has not been made at all. In this sense, the game falls into the general class of dynamic strategic models with imperfect private monitoring. More specifically, it is a stochastic game with privately observed states, where a player's state is the last action picked and not yet implemented.

The parameter β_i captures player i's "agility". A more agile player is represented by a higher value of β. The interpretation is that players are not individuals but *organizations*. One of the characteristics of a large, heavy organization is that it takes time from the moment a decision to change a course of action is made until the moment it is finally implemented. My objective is to explore the implications of such "heaviness" for repeated interaction between organizations. I assume that $\beta_1 \leq \beta_2$. The interpretation is that agility and strength are negatively correlated. Thus, when both organizations choose the same arena at period t (i.e., $a_1^t = a_2^t$), the heavier organization will beat the more agile organization because it is stronger. When the two organizations choose different arenas (i.e., $a_1^t \neq a_2^t$), the conflict is avoided and thus the outcome is a draw. Player 2's superior agility enables him to avoid head-to-head conflict with his stronger, heavier rival.

The model abstracts from other aspects of limited agility. For example, I stick to the conventional assumption that monitoring the opponent's action is quick, such that every switch in his behavior is immediately detected (unlike the works cited in footnote 1). In addition, the assumption that players are unable to make any choice while their decision to switch is waiting to be implemented means that they cannot issue a retraction. One could imagine a model in which the organization chooses to reverse a decision to switch while it is still waiting to be implemented, and retraction also takes time. These examples demonstrate that once we open the door for delays in the decision process, many interesting factors could be integrated into the repeated-game model. I hope to address these aspects in future work.

3 Analysis

The analysis in this section holds for an arbitrary discount factor. Recall that when $\beta_1 = \beta_2 = 1$, we are back with standard repeated Matching Pennies. In this case, there is a unique Nash equilibrium, in which each player chooses each action with equal probability at every history. The Minimax Theorem holds, such that this is also the unique max-min strategy for each player.

At any period-t history h^t in which player i gets to act, let $k(i, h^t)$ denote the number of periods that elapsed since his latest observed switched action. That is, player i gets to act at period t, and $a_i^{t-1} = \cdots = a_i^{t-1-k(i,h^t)} \neq a_i^{t-2-k(i,h^t)}$. In particular, if $a_i^{t-1} \neq a_i^{t-2}$, then $k(i, h^t) = 0$.

Proposition 1 *There is a unique Nash equilibrium. At any history h in which player i gets to act, he chooses to switch his action with probability $\frac{1}{2}\beta^{k(i,h)-2}$.*

Proof. The method of proof is as follows. I first show that each player has a unique strategy that induces an actual (i.e., observed) switching probability of $\frac{1}{2}$ after every history. Using the Minimax Theorem, this will imply that this must be the player's unique Nash equilibrium strategy.

Let $p_i(k)$ be the probability that player i decides to switch, conditional on the event that k periods elapsed since his last switch and that he chose not to switch in each one of them. Then, for any k, the probability that the player chooses not to switch at all periods $1, ..., k$ since the last switch is

$$\prod_{k'=1}^{k} (1 - p_i(k'))$$

The probability that the player does not actually switch at all periods $1, ..., k$ is

$$\prod_{k'=1}^{k} (1 - \beta_i p_i(k'))$$

Thus, the probability that player i chose not to switch at all periods $1, ..., k$ conditional on the observation that he did not actually switch during those periods is

$$x_i(t) = \frac{\prod_{k'=1}^{k} (1 - p_i(k'))}{\prod_{k'=1}^{k} (1 - \beta_i p_i(k'))}$$

Note that by definition, this function weakly decreases in k. Our objective is to construct a function $p_i(k)$ such that

$$x_i(k-1) \cdot \beta_i p_i(k) + (1 - x_i(k-1)) \cdot \beta_i = \frac{1}{2}$$

because the L.H.S of this equation is the probability that there is an observed switch at period k conditional on having no observed switch in periods $1, ..., k-1$. If we can find a function that satisfies this equation for every k, then we will have generated a stationary observed switching rate of $\frac{1}{2}$. Note that $x_i(0) = 0$ by definition. Thus, we can solve this equation by induction and obtain the unique solution

$$\begin{aligned} p_i(k) &= \frac{1}{2}(\beta_i)^{k-2} \\ x_i(k) &= \frac{2\beta_i - 1}{\beta_i(2 - (\beta_i)^{k-1})} \end{aligned}$$

The function $x_i(k)$ is strictly decreasing, and asymptotes at $1 - \frac{1}{2\beta_i}$ when $k \to \infty$.

A stationary observed switching rate of $\frac{1}{2}$ coincides with the max-min strategy in the case of $\beta_1 = \beta_2 = 1$. The set of feasible strategies available to player i when $\beta_i < 1$ is contained in the feasible set for the benchmark case of $\beta_i = 1$, because when $\beta_i < 1$

there are histories in which he is unable to act and restricted to a fixed switching rate of β_i. Therefore, the max-min payoff for each player is weakly below the benchmark case. Since each player i can generate the same realized switching rate as in the benchmark, using a uniquely defined strategy, by the Minimax Theorem it follows that this is also the unique Nash equilibrium strategy for the player. ■

Thus, players' equilibrium payoff is exactly as in the benchmark case with unlimited agility. In other words, the heavy player 1 is not harmed by his inferior agility. The players' observed behavior is indistinguishable from the benchmark, too. (These features would break down if we relaxed the assumption that $\beta_1, \beta_2 \geq \frac{1}{2}$.) However, the players' behavior differs markedly from the benchmark. Players become less likely to choose to switch as time goes on, and in the limit they are completely inert. In other words, players display hysteresis. Moreover, the less agile the player, the greater the tendency for hysteresis. This is consistent with the intuition that stronger, heavier organizations are also slower to change their behavior. In other words, greater *exogenous* slowness in *implementing* changes leaders to greater *endogenous* slowness in *initiating* them.

References

Abreu, D. and A. Rubinstein (1988), "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," Econometrica 56, 1259-1282.

Bhaskar, V. and I. Obara (2011), "Firm Reputation with Information Lags," mimeo.

Fudenberg, D., Y. Ishii and S. Kominers (2014), "Delayed-Response Strategies in Repeated Games with Observation Lags," Journal of Economic Theory 150, 487-514.

Lipman, B. and R. Wang (2000), "Switching Costs in Frequently Repeated Games," Journal of Economic Theory 93, 149-190.

Rubinstein, A. (1986), "Finite Automata Play the Repeated Prisoner's Dilemma," Journal of Economic Theory 39, 83-96.

Scarf, H. and L. Shapley (1957), "Games with Partial Information," Contributions to the Theory of Games 3, 213-229.