The Regularity of the Value Function of Repeated Games with Switching Costs

Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Regularity of the Value Function of Repeated Games with Switching Costs

Yevgeny Tsodikovich, Xavier Venel, Anna Zseleva

informs
https://pubsonline.informs.org/journal/moor

MATHEMATICS OF OPERATIONS RESEARCH
Articles in Advance, pp. 1–7
ISSN 0364-765X (print), ISSN 1526-5471 (online)

The Regularity of the Value Function of Repeated Games with Switching Costs

Yevgeny Tsodikovich,[a,b,*] **Xavier Venel,**[c] **Anna Zseleva**[d]

[a] Aix-Marseille School of Economics, French National Centre for Scientific Research (CNRS), Aix Marseille University, 13205 Marseille, France; [b] Department of Economics, Faculty of Social Sciences, Bar-Ilan University, 5290002 Ramat-Gan, Israel; [c] Dipartimento di Economia e Finanza, Luiss University, 00197 Rome, Italy; [d] Department of Quantitative Economics, School of Business and Economics, Maastricht University, 6200 MD Maastricht, Netherlands

*Corresponding author

Contact: yevgets@gmail.com, https://orcid.org/0000-0001-7643-4073 (YT); xvenel@luiss.it, https://orcid.org/0000-0003-1150-9139 (XV); anna.zseleva@maastrichtuniversity.nl, https://orcid.org/0000-0001-5530-1183 (AZ)

Received: November 15, 2021
Revised: May 13, 2022; September 2, 2022
Accepted: September 17, 2022
Published Online in Articles in Advance: November 2, 2022

MSC2020 Subject Classification: Primary: 91A15, 91A05, 91A10

https://doi.org/10.1287/moor.2022.1325

Abstract. We study repeated zero-sum games where one of the players pays a certain cost each time he changes his action. We derive the properties of the value and optimal strategies as a function of the ratio between the switching costs and the stage payoffs. In particular, the strategies exhibit a robustness property and typically do not change with a small perturbation of this ratio. Our analysis extends partially to the case where the players are limited to simpler strategies that are history independent-namely, static strategies. In this case, we also characterize the (minimax) value and the strategies for obtaining it.

Funding: The project leading to this publication has received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency [Grant ANR-17-EURE-0020] and from the Excellence Initiative of Aix-Marseille University–A*MIDEX. Y. Tsodikovich was supported in part by the Israel Science Foundation [Grants 2566/20, 1626/18, and 448/22]. X. Venel acknowledges the financial support of the French National Research Agency through Project CIGNE [ANR-15-CE38-0007-01].

Keywords: switching costs • repeated games • stochastic games • zero-sum games

1. Introduction

We consider a repeated normal-form zero-sum game where at each time step, the minimizing player pays the maximizer both the "standard" outcome of the game and an additional fine if he switched his previous action. This model—namely, the *traveling inspector model*—was presented in Filar and Schultz [4] and is used to study different scenarios with switching costs (Darlington et al. [1], Xu et al. [14], and Yavuz and Jeffcoat [15], to mention a few).

A game with switching costs is equivalent to a stochastic game where the states correspond to the previous action taken by the minimizing player (Filar and Schultz [4]). Because only the minimizing player controls the states, this is a single-controller stochastic game, and there exists an optimal strategy that is *stationary* (depends solely on the state) and can be computed using one of several standard tools, such as those in Raghavan and Syed [9] and Raghavan [8]. We characterize completely the value and the optimal stationary strategies, and we show that they belong to a finite set that depends only on the underlying one-shot game, not on the switching costs. Our main finding is that the value function is piecewise linear in the weight between the switching costs and the "standard" stage payoff, c. Consequently, the optimal strategy depends only on the segment in which c is situated, not its value, and the minimizer can play optimally without knowing the exact weight (as in Rass and Rainer [10]).

In some applications there is an additional requirement to use only history- and time-independent strategies (denoted as *static strategies*)-that is, to use the same mixed action in every stage (Rass et al. [11], Schoenmakers et al. [12]). We repeat our analysis with this constraint and study the properties of the optimal payoff[1] in static strategies and the corresponding optimal strategies. In general, the optimal payoff is not piecewise linear, but the previous result for stationary strategies carries on in the special case of switching costs that are independent of the actions. In this case, our results theoretically explain why in empirical works such as Liuzzi et al. [7] the optimal static strategies only change slightly (if at all) in response to a small change in the ratio between the switching costs and the stage payoffs.

2. The Switching Costs Model

A *zero-sum game with switching costs* is a tuple $\Gamma = (A, S, c)$, where $A = (a_{ij})$ is an $m \times n$ matrix, $S = (s_{ij})$ is an $n \times n$ matrix with nonnegative entries, and $c \geqslant 0$. At time step t, player 1 (the maximizer) chooses an integer $i(t)$ in the set $[m] = \{1, \ldots, m\}$, and player 2 (the minimizer) chooses $j(t) \in [n]$. The stage payoff that player 2 pays player 1 is $a_{i(t)j(t)} + cs_{j(t-1)j(t)}$, where $cs_{j(t-1)j(t)}$ represents the cost of switching.[2] We assume that keeping the same action is costless ($s_{jj} = 0$). The extension to mixed strategies is standard, but it should be noted that $j(t-1)$ is known at time t, even if it was determined by a mixed action.

The process repeats indefinitely, and the payoff is the undiscounted average. More precisely, let (σ, τ) be a pair of strategies in the repeated game, and denote by $i(t), j(t)$ the chosen actions at time t according to σ, τ (given the history). We set $j(0) = 1$ (an assumption that has no effect on the payoff) and define the payoff to be

$$\gamma(\sigma, \tau) = \liminf_{T \to \infty} \mathbb{E}_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} a_{i(t)j(t)} + cs_{j(t-1)j(t)}\right).$$

This game is equivalent to a stochastic game where each state represents the previous action of player 2. The sets of actions in all states are $[m]$ and $[n]$, the payoff in state k when player 1 plays i and player 2 plays j is $a_{ij} + cs_{kj}$, and the next state is j. The class of games where only one player controls the state transitions was studied by Filar [2], who showed that the value exists and is obtained in strategies that depend solely on the current state, not on the history-namely, *stationary strategies*. It follows that the value of Γ exists, and it is obtained in strategies that depend solely on the previous action of player 2.

Definition 1. A *stationary strategy* is a strategy that depends in each t on the previous action of player 2, $j(t-1)$, but not on the rest of the history or t itself. Hence, a stationary strategy is a vector of n mixed actions, one to follow each possible pure action of player 2.

Definition 1 concerns both players, but in any case, the dependence is on the previous action of player 2, as he is the one paying switching costs and controlling the states.

Following the literature (Liuzzi et al. [7], Rass and Rainer [10], Schoenmakers et al. [12]), we also consider strategies independent of the time and of the history.

Definition 2. A *static strategy* is a strategy that plays the same mixed action in each stage, regardless of the history. Hence, a static strategy is one mixed action played repeatedly.

Suppose player 1 uses the static strategy x and player 2 uses the static strategy y. We obtain a closed-form formula for the payoff in matrix notation:

$$g(c)(x, y) = x^T Ay + cy^T Sy. \tag{1}$$

Typically, the value may not exist in static strategies. Instead, the figure of merit studied in the literature is the minimax, defined as $\tilde{v}(c) = \min_y \max_x g(c)(x, y)$. For simplicity, we refer to it as the *value in static strategies*, but it should be understood only as the *minimax value*.

3. Results

3.1. A Useful Lemma for Parametric One-Shot Games

We study the value of a parametric one-shot game, where the payoff is linear in each parameter in a way that is independent of the actions of player 1. The lemma stands alone, as it might be of general interest. Formally, let A be an $m \times n$ zero-sum game, and let $b_1, \ldots, b_n$ be nonnegative constants. For each $\underline{x} = (x_1, \ldots, x_n)$, the game $\Gamma(\underline{x})$ is defined to be the one-shot zero-sum game whose payoff matrix is $a_{ij} + b_j x_j$, and its value is denoted by $v(\underline{x})$.

Lemma 1. *The value function $v(\underline{x}) : \mathbb{R}^n \to \mathbb{R}$ is continuous, increasing in every parameter, concave, and piecewise linear in every direction.*

Proof. Let I_1 and I_2 be some subsets of the rows and columns, respectively, of the one-shot game A. We check whether player 2 can make player 1 indifferent among all the actions in I_1 using a completely mixed action over I_2 and whether player 1 prefers them over the actions not in I_1. Fix $k \in I_1$. We look for a vector $y \in \mathbb{R}^n$ of player 2 such that the support of y is exactly I_2 ($\sum y_j = 1$ with $\forall j \in I_2 : y_j > 0$, and $\forall j \notin I_2 : y_j = 0$), player 1 is indifferent among the actions in I_1 and prefers actions in I_1 over other actions:

$$\begin{cases} \sum_{j \in I_2} y_j a_{kj} = \sum_{j \in I_2} y_j a_{lj} & \forall l \in I_1 \backslash \{k\}, \\ \sum_{j \in I_2} y_j a_{kj} \geqslant \sum_{j \in I_2} y_j a_{lj} & \forall l \notin I_1. \end{cases} \tag{2}$$

If there exists solutions to this system of equations, we denote one of them by $y(I_1, I_2)$.

Suppose player 2 uses $y(I_1, I_2)$ in $\Gamma(\underline{x})$ for some $\underline{x}$. Player 1 is indifferent among the actions in I_1 and prefers them over other actions if Equation (2) holds with $(a_{\cdot j} + b_j x_j)$ instead of $a_{\cdot j}$. This is indeed the case, as the $b_j x_j$ terms cancel out. Therefore, when player 1 best responds, he chooses an action in $\Delta(I_1)$, and the variable x_j appears in the payoff linearly with the slope $y_j b_j$. Denote by $l(I_1, I_2)(\underline{x})$ the payoff function, which is linear in $\underline{x}$.

Fix $\underline{x}$ and a pair of optimal strategies (p^*, q^*). This profile has a corresponding support pair $(I_1(\underline{x}), I_2(\underline{x}))$. By definition, $y(I_1(\underline{x}), I_2(\underline{x}))$ makes player 1 indifferent between the actions he chooses with nonzero probability according to p^*; hence the pair $(p^*, y(I_1(\underline{x}), I_2(\underline{x})))$ is optimal and so does $y(I_1(\underline{x}), I_2(\underline{x}))$. It follows that $v(\underline{x}) = l(I_1(\underline{x}), I_2(\underline{x}))(\underline{x})$. Moreover, the value function $v(\underline{x})$ is continuous in each x_i (this is a polynomial game; Shapley and Snow [13]); hence the support pairs can only change when the payoffs they induce are equal. Because there is a finite number of supports, there is a partition of $\mathbb{R}^n$ into finitely many disjoint convex and closed sets (one for each pair of supports) such that the restriction of $v(\underline{x})$ to each of them is the restriction of some linear map to this set. We conclude that $v(\underline{x})$ is piecewise linear in any direction of $\mathbb{R}^n$, and because the slopes are convex combinations of the nonnegative coefficients b_i, $v(\underline{x})$ is also increasing in each variable.

Finally, $v(\underline{x})$ is concave. Let $\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3 \in \mathbb{R}^n$ be three points on the same line such that $v(\underline{x})$ is linear on the segments $[\boldsymbol{c}_1, \boldsymbol{c}_2]$ and $[\boldsymbol{c}_2, \boldsymbol{c}_3]$. For $\boldsymbol{c}^* \in (\boldsymbol{c}_1, \boldsymbol{c}_2)$, let $y(\boldsymbol{c}^*)$ be the optimal action of player 2 chosen using the above-mentioned method. If player 2 plays $y(\boldsymbol{c}^*)$ regardless of $\boldsymbol{c}$, the payoff is a linear function of $\boldsymbol{c}$ that coincides with the value on $[\boldsymbol{c}_1, \boldsymbol{c}_2]$. In the region $[\boldsymbol{c}_1, \boldsymbol{c}_2]$, the value function must be below this line, so the slope of $\boldsymbol{c}$ must decrease. To conclude, on this line, $v(\underline{x})$ is a piecewise linear function with decreasing slope (i.e., concave). □

3.2. The Main Results

3.2.1. The Value Function in Stationary Strategies. Our first main result is the following theorem.

Theorem 1. *For every $c \geqslant 0$, the game has a value in stationary strategies denoted by $v(c)$. This function is continuous, increasing, concave, piecewise linear, and eventually constant. If $v(c)$ is linear on $[\underline{c}, \overline{c}]$, then player 2 has a strategy that is optimal for all $c \in [\underline{c}, \overline{c}]$.*

Proof. This game is equivalent to a stochastic game where the current state corresponds to the pure action chosen by player 2 in the previous time period, so the set of states is also $[n]$. In this stochastic game, only one player controls the transitions, so the existence of the value follows from the existence of the value in Filar [2]. Moreover, according to Filar and Raghavan [3], the value of the game is the same as the value of a one-shot normal-form game, whose pure actions are the pure stationary actions in the stochastic game (each pure action in the one-shot game is a vector of size n dictating which pure action to choose at each state). Given a pure strategy of player 1 and a pure strategy of player 2, the payoff is linear in c with coefficients depending on the strategies. Moreover, the coefficient of c is determined only by the strategy of player 2 because only he bears switching costs. Hence, there exists a_{ij} and b_j such that the payoff can be denoted by $a_{ij} + b_j c$, with $b_j \geqslant 0$ because it is a convex combination of the elements of S.

Lemma 1 can be applied to this one-shot game by setting for all $i: x_i = c$, so $v(c) = v(c, \ldots, c)$, and it is continuous, concave, increasing, and piecewise linear. Because $v(c)$ is bounded by the pure minimax of A, it cannot strictly increase for all c and is eventually constant.

Suppose $v(c)$ is linear on $[\underline{c}, \overline{c}]$, and fix $c^* \in (\underline{c}, \overline{c})$. Let σ_{c^*} be an optimal strategy for player 2 in the corresponding stochastic game, and suppose player 2 plays σ_{c^*} regardless of c, whereas player 1 best responds (as a function of c). When player 2 plays σ_{c^*}, he fixes the transition probabilities between states, and these are now independent of c. Hence, the fraction of time spent in each state is constant and independent of c. In state k, the payoff matrix is of the form $a_{ij} + s_{kj} c$, so the part corresponding to c in each column is identical and linear with the same slope, regardless of the row. When mixing the columns in state k according to σ_{c^*}, the expected payoff of the rows are parallel lines. Player 1 best responds in each state, but because the lines are parallel, his choice does not depend on c. Thus, the payoff in each state is a linear function of c, and the total payoff, which is a weighted average of these functions with the percentage of time spent in each state as weights, is a linear function of c, denoted by $f_{c^*}(c)$. Because $v(c)$ is the value, $v(c) \leqslant f_{c^*}(c)$, with equality at c^*. These two lines that intersect once must coincide, and σ_{c^*} obtains the value in the entire segment. □

Optimal strategies are therefore robust to small changes of c: knowing the exact c is not necessary to play optimally, and it is almost universally unnecessary to adjust the strategy as c changes. Moreover, if player 2 wishes to minimize separately his stage payoff in the repeated game and his switching costs, an alternative approach is to consider a game where the payoff is their convex combination, as in our model (Rass and Rainer [10]). Theorem 1 shows that the exact weights of the two goal functions are of small significance for player 2, whose optimal strategy comes from a finite set that depends solely on A. Interestingly, this is not true for player 1, and his optimal strategy typically depends on the exact c.

The concavity of $v(c)$ comes from a compromise between the stage payoff and the switching costs. The higher the c, the more costly it is to switch, so it is better to play a strategy that rarely changes actions in expense of some loss in the stage game A.

To conclude, the family of functions described in Theorem 1 is the widest possible. Any function that holds these properties corresponds to a zero-sum game with switching costs; hence our theorem provides a complete characterization of the value function of such games.

Remark 1. Let $v(c): [0,\infty) \to \mathbb{R}$ be a continuous, increasing, concave, piecewise linear function and eventually constant. Then, there exists a one-shot game A and a switching costs matrix S such that $v(c)$ is the value of the game (A, S, c).

Proof. We now provide a sketch of the proof of Remark 1. The value function of the game

$$A = \begin{pmatrix} 2b_1 & 0 & 2b_2 & 0 & \overline{v} \\ 0 & 2b_1 & 0 & 2b_2 & \overline{v} \end{pmatrix}$$

with

$$S = \begin{pmatrix} 0 & 2\beta_1 & M & M & M \\ 2\beta_1 & 0 & M & M & M \\ M & M & 0 & 2\beta_2 & M \\ M & M & 2\beta_2 & 0 & M \\ M & M & M & M & 0 \end{pmatrix},$$

where $M \gg \beta_i$, is

$$v(c) = \begin{cases} b_1 + \beta_1 c & \text{if } c \in [0, c_1], \\ b_2 + \beta_2 c & \text{if } c \in [c_1, c_2], \\ \overline{v} & \text{if } c \in [c_2, \infty). \end{cases}$$

The generalization to $v(c)$ with more than three segments is straightforward. □

3.2.2. The Value Function in Static Strategies. The minimax value in static strategies exhibits similar properties to $v(c)$ except piecewise linearity. Piecewise linearity can be obtained by adding the common assumption that the switching costs are independent of the switched actions (i.e., $s_{ij} = 1$ for $i \neq j$). This case arises when the costs stem from the act of "switching" itself and do not depend on the actions being switched (Lipman and Wang [5, 6], Schoenmakers et al. [12]).

Theorem 2. *The minimax value $\tilde{v}(c)$ in static strategies is a continuous, increasing, and concave semialgebraic function.*

If, in addition, the switching costs are uniform ($s_{ij} = 1$ for all $i \neq j$), then $\tilde{v}(c)$ is piecewise linear, and for each $[\underline{c}, \overline{c}]$ where $\tilde{v}(c)$ is linear, there exists a static strategy for player 2 that is optimal for all $c \in [\underline{c}, \overline{c}]$.

Proof. Consider $g(c)(x,y)$ from Equation (1) and recall that, by definition,

$$\tilde{v}(c) = \min_y \max_x g(c)(x,y) = \min_{y \in \Delta([n])} \left\{ \max_{x \in \Delta([m])} \{x^T A y\} + c y^T S y \right\}. \tag{3}$$

Note that arg $\max_x g(c)(x,y)$ depends solely on y and not on c. Because $g(c)(x,y)$ is continuous, so too is $\tilde{v}(c)$. Moreover, $\tilde{v}(c)$ is semialgebraic because $g(c)(x,y)$ is a polynomial in each variable.

We next prove that the function $\tilde{v}(c)$ is increasing. Let $0 \leqslant c_1 < c_2$, and let y_{c_2} be the arg min from Equation (3). Here, $\tilde{v}(c_2) = \max_{x \in \Delta([m])}\{g(c_2)(x, y_{c_2})\} \geqslant \max_{x' \in \Delta([m])}\{g(c_1)(x', y_{c_2})\} \geqslant \tilde{v}(c_1)$, where the first inequality follows from $c_1 < c_2$ and $S \geqslant 0$ (the maximizing x is the same, as it depends solely on $\boldsymbol{y}_{c_2}$). The last inequality follows from the definition of $\tilde{v}(c_1)$.

To prove that the function $\tilde{v}(c)$ is concave, first let $\beta \in (0,1)$. Then

$$\begin{aligned}\tilde{v}(\beta c_1+(1-\beta)c_2) &= \min_y \left\{ \max_x \{x^T Ay\} + (\beta c_1 + (1-\beta)c_2) y^T Sy \right\} \\ &\geqslant \min_y \left\{ \beta \max_x g(c_1)(x,y) \right\} + \min_y \left\{ (1-\beta) \max_x g(c_2)(x,y) \right\} \\ &= \beta \tilde{v}(c_1) + (1-\beta)\tilde{v}(c_2).\end{aligned}$$

The inequality is obtained because term-by-term minimization yields a smaller result than minimizing the entire sum. This completes the main part of the proof.

In addition, suppose that $s_{ij} = 1$ *for* $i \neq j$. Under this assumption, $y^t Sy = 1 - \|y\|^2$, and $\tilde{v}(c) = c + \min_{y \in \Delta([n])} \left\{ \max_{x \in \Delta([m])} \{x^T Ay\} - c\|y\|^2 \right\}$. For each $i \in [m]$, let I_i be the subset of $\Delta([n])$ such that i is a best response for player 1 in the one-shot game A to all $y \in I_i$. It is well known that I_i is a compact, convex polyhedron. Now, suppose player 2 is restricted to playing only in I_i. The best response for player 1 is i, and the payoff is $c + \min_{y \in I_i} \left\{ A_i y - c\|y\|^2 \right\}$, where A_i is the ith row of A. This is a minimization of a concave function over a polyhedron, so the minimum is attained at one of its vertices. There are m polyhedrons, and each one has finitely many vertices, so the set of candidate optimal strategies is finite. Because for a given pair of strategies the payoff is linear in c, the value function for each c coincides with one of the finitely many linear functions (one for each candidate for being the optimal strategy). Because the value function is also continuous, it can change the linear function it coincides with only when the two linear functions intersect, which implies that $\tilde{v}(c)$ is piecewise linear, and on each segment where it is linear, there exists a static strategy that obtains the value for all c's in the segment. □

The requirement for uniform switching costs is essential for the piecewise linearity of $\tilde{v}(c)$. For example, it is easy to verify that for the game

$$A = \begin{pmatrix} -200 & 1 & 200 \\ 200 & 1 & -200 \end{pmatrix}$$

with the switching costs matrix

$$S = \begin{pmatrix} 0 & 1 & 100 \\ 1 & 0 & 1 \\ 100 & 1 & 0 \end{pmatrix},$$

the optimal static strategy for player 2 strongly depends on c, and in the domain $c \in \left[\frac{1}{98}, \frac{1}{2}\right]$, the value is $\tilde{v}(c) = 1 - \frac{(1-2c)^2}{192c}$.

With the additional assumption that the switching costs are uniform, the optimal static strategy comes from a finite set and is robust to the exact value of c. This is also true for player 1 (unlike in the case of stationary strategies), because in this case, player 1 best responds to the static strategy of player 2, and he can always choose a pure best reply. This provides a theoretical explanation for the numerical results of Rass and Rainer [10] and Liuzzi et al. [7], which indicate that slightly changing c has a small effect, if any, on the optimal strategies.

This result can be used to search more efficiently for the optimal strategy within the set, by eliminating from consideration strategies with too-high expected per-stage switching cost. The idea is that if a particular strategy y_{c^*} is optimal for some c^*, the concavity dictates that y is not optimal for $c > c^*$ if $y^T Sy > y_{c^*}^T Sy_{c^*}$. A similar approach can be used for the optimal stationary strategy as a consequence of Theorem 1, although calculating the slope for a given stationary strategy is more complicated.

3.2.3. 2 × 2 Games. We finalize by showing that in 2 × 2 games, the value and the minimax value in static strategies coincide. This generalizes theorems 4.1 and 4.2 and corollary 4.1 in Schoenmakers et al. [12] for the case that $s_{12} \neq s_{21}$. The main idea is that there are only two candidates for being the optimal stationary strategy: the optimal strategy of A and the pure minimax of A; both are also static strategies.

Proposition 1. *Let*

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

be a zero-sum game, and let

$$S = \begin{pmatrix} 0 & s_{12} \\ s_{21} & 0 \end{pmatrix}$$

be the switching costs of player 2. *The value of* (A, S, c) *can be obtained by a static strategy* (*i.e.*, $\tilde{v}(c) = v(c)$).

Proof. If the value of A without switching costs can be obtained in pure actions, then $v(c) = \tilde{v}(c) = v$ for all c. Otherwise, without loss of generality, $\alpha > \beta, \gamma$ and $\delta > \beta, \gamma$. Let $p \in (0,1)$ be the unique probability of playing the left column (action L) in A that achieves the value.

In each state, in a similar way to Equation (4), player 1 is indifferent between his two actions if and only if player 2 plays p; otherwise, player 1 plays purely one of them, regardless of the switching costs. Thus, in each state, either player 2 plays purely too or he plays the mixed action p. There remains only four types of possible stationary strategies to consider: (i) play p in both states, (ii) play in a way that never reaches one of the states (e.g., play L after s_L), (iii) play purely but always visit all the states (play L in state s_R and R in s_L), and (iv) play purely in only one state and p in the other (e.g., R in state S_L).

Options (i) and (ii) correspond to static strategies, whereas option (iii) is clearly not optimal. We show by contradiction that a strategy of type (iv) is not optimal. Let κ be the continuation payoff after playing R (the continuation payoff after playing L is normalized to 0). When we add the switching costs and the continuation payoffs to the games, we obtain two one-shot games:

$$s_L : \begin{pmatrix} \alpha & \beta + \kappa + cs_{12} \\ \gamma & \delta + \kappa + cs_{12} \end{pmatrix} \qquad s_R : \begin{pmatrix} \alpha + cs_{21} & \beta + \kappa \\ \gamma + cs_{21} & \delta + \kappa \end{pmatrix}. \tag{4}$$

To be optimal in the stochastic game, the profile must be optimal in each of these one-shot games. Consider state s_L. Because it is optimal to play purely R, we necessarily have $\alpha \geqslant \beta + \kappa + cs_{12}$ or $\gamma \geqslant \delta + \kappa + cs_{12}$. If only the first equation is true, the equilibrium is mixed, and R is not optimal. If only the second equation is true, then there is a contradiction: $\alpha > \gamma \geqslant \delta + \kappa + cs > \beta + \kappa + cs > \alpha$. It follows that both equations are true—that is, the right column dominates the left—and in particular, $\gamma \geqslant \delta + cs_{12} + \kappa$.

In state s_R, however, the optimal strategy is mixed, and therefore such dominance is impossible. If the direction of the original inequalities (on $\alpha, \beta, \gamma, \delta$) remains correct, then $\delta + \kappa \geqslant \gamma + cs_{21}$ and $s_{12} + s_{21} \leqslant 0$, which is a contradiction. In the other case that both inequalities change direction, the contradiction is constructed using the first row. □

Acknowledgments

The authors thank E. Solan, D. Lagziel, J. Flesch, and G. Ashkenazi-Golan, as well as the editor and the anonymous referee for their highly valuable comments. X. Venel is a member of GNAMPA-INdAM (Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni).

Endnotes

[1] Typically, the game has no value in static strategies. Hereafter, when discussing the value in such strategies, we are referring to the minimax value and when discussing optimal strategies, the minimax strategies.

[2] We emphasize that the game remains a zero-sum game, as the switching costs are transferred to the adversary, player 1. If part of the switching costs dissipate, the game is no longer a zero-sum game.

References

[1] Darlington M, Glazebrook KD, Leslie DS, Shone R, Szechtman R (2022) A stochastic game framework for patrolling a border. Preprint, submitted May 20, https://doi.org/10.48550/arXiv.2205.10017.

[2] Filar J (1980) Algorithms for solving some undiscounted stochastic games. PhD thesis, University of Illinois at Chicago Circle, Chicago.

[3] Filar J, Raghavan T (1984) A matrix game solution of the single-controller stochastic game. *Math. Oper. Res.* 9(3):356–362.

[4] Filar J, Schultz T (1986) The traveling inspector model. *Oper.-Res.-Spektrum* 8(1):33–36.

[5] Lipman B, Wang R (2000) Switching costs in frequently repeated games. *J. Econom. Theory* 93(2):149–190.

[6] Lipman B, Wang R (2009) Switching costs in infinitely repeated games. *Games Econom. Behav.* 66(1):292–314.

[7] Liuzzi G, Locatelli M, Piccialli V, Rass S (2021) Computing mixed strategies equilibria in presence of switching costs by the solution of nonconvex QP problems. *Comput. Optim. Appl.* 79(3):561–599.

[8] Raghavan TES (2003) Finite-step algorithms for single-controller and perfect information stochastic games. Neyman A, Sorin S, eds. *Stochastic Games and Applications* (Springer, Dordrecht, Netherlands), 227–251.

[9] Raghavan TES, Syed Z (2002) Computing stationary Nash equilibria of undiscounted single-controller stochastic games. *Math. Oper. Res.* 27(2):384–400.

[10] Rass S, Rainer B (2014) Numerical computation of multi-goal security strategies. Poovendran R, Saad W, eds. *Proc. 5th Internat. Conf. Decision Game Theory Security* (Springer, Cham, Switzerland), 118–133.
[11] Rass S, König S, Schauer S (2017) Defending against advanced persistent threats using game-theory. *PLoS One* 12(1):e0168675.
[12] Schoenmakers G, Flesch J, Thuijsman F, Vrieze O (2008) Repeated games with bonuses. *J. Optim. Theory Appl.* 136(3):459–473.
[13] Shapley L, Snow R (1950) Basic solutions of discrete games. Kuhn HW, Tucker AW, eds. *Contributions to the Theory of Games*, Vol. 1, Annals of Mathematics Studies, No. 24 (Princeton University Press, Princeton, NJ), 27–36.
[14] Xu H, Ford B, Fang F, Dilkina B, Plumptre A, Tambe M, Driciru M, et al. (2017) Optimal patrol planning for green security games with black-box attackers. *Proc. 8th Internat. Conf. Decision Game Theory Security* (Springer, Cham, Switzerland), 458–477.
[15] Yavuz M, Jeffcoat D (2007) An analysis and solution of the sensor scheduling problem. Pardalos PM, Murphey R, Grundel D, Hirsch MJ, eds. *Advances in Cooperative Control and Optimization* (Springer, Berlin), 167–177.