



The Price of History-Independent Strategies in Games with Inter-Temporal Externalities

Yevgeny Tsodikovich¹ · Xavier Venel² · Anna Zseleva³

Accepted: 24 January 2024 / Published online: 2 March 2024
© The Author(s) 2024

Abstract

In this paper, we compare the value of zero-sum stochastic games under optimal strategies (that are, for single-controller stochastic games, stationary) to the commonly used time-independent strategies (“static strategies”). Our findings are summarized in a series of theorems which provide the lower bound on the optimality of the static strategy under different assumptions. These bounds can be used to assess whether the additional computational complexity is worth the extra payoff gain or, symmetrically, assess the price of playing sub-optimal but simple strategies when stationary ones are forbidden.

Keywords Switching costs · Repeated games · Stochastic games · Zero-sum games

Mathematics Subject Classification C72 · C73

1 Introduction

Consider the *Traveling Inspector Model* (TIM, Filar and Schultz [5]) in which an inspector patrols between different possible illegal dump sites while paying some movement costs. A natural way to model this game is as a stochastic game and it is known that both the optimal

The authors thank E. Solan, D. Lagziel, as well as the editor and the two anonymous referees for their highly valuable comments.

✉ Yevgeny Tsodikovich
yevgets@gmail.com

Xavier Venel
xvenel@luiss.it

Anna Zseleva
anna.zseleva@maastrichtuniversity.nl

¹ Department of Economics, Bar Ilan University, Ramat Gan, Israel

² Dipartimento di Economia e Finanza, LUISS, Viale Romania 32, 00197 Rome, Italy

³ School of Business and Economics, Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 Maastricht, MD, The Netherlands

patrolling strategy of the inspector and the optimal polluting strategy of the perpetrators depend on the inspector's current location [6]. They are stationary.

Several papers in the computer science literature have argued in favor of looking at simpler strategies that are time-independent. Rass et al. [14] argue that “if a player would implement an equilibrium mixed strategy by switching such that the costs for playing the equilibrium strategy is minimized, then the opponent can gain quite some idea what the next move of the defender will be” and deduce that the analysis should focus on history-independent strategies (“static strategies”). This statement is also sustained in Wachter et al. [18] and Liuzzi et al. [9]. These articles focus then on the computational aspects of the optimal static strategy or an approximation of it. Independently [15] introduced them under the name of “simple strategies” and studied when they are optimal.

Furthermore, Rass et al. [14] and Wachter et al. [18] argue that optimal strategies are too complex to be computed or implemented for real-life scenarios. Nevertheless, finding the optimal static strategy is not straightforward and can even be an NP-hard problem [9].

The aim of this article is to contribute to this discussion of simple strategies versus optimal ones by comparing the payoffs guaranteed by both types of strategies. We choose to adopt a theoretical approach starting with a generic structure of stochastic games to narrow down to more and more specific structures. With the right structures, we can show that static strategies guarantee a non-trivial fraction of the optimal payoff.

We first notice that without a specific structure, static strategies can be arbitrarily bad (Example 2). As a consequence, we limit the discussion to single-controller stochastic games¹ with deterministic and state-independent transitions, and show that at least quarter of the value can be guaranteed with a static strategy.

We then narrow the scope to stochastic games that are equivalent to repeated zero-sum games with an arbitrary structure of inter-temporal externalities, including switching costs as in the TIM or bonuses as in [15]. We show that half of the value can be obtained with static strategies when the externalities are non-negative. With more assumptions on the structure of the inter-temporal externalities, we can provide several additional specific bounds in particular when players are penalized for switching and the cost is independent of the actions (as in [7, 8, 15]). To the best of our knowledge, such a study of the “price of being static” has never been conducted before.

Structure of the paper. The model for general stochastic games is presented in Sect. 2. The main result for single-controller stochastic games with deterministic and state-independent transitions is presented in Sect. 3. Section 4 presents the results for stochastic games that are equivalent to repeated games with inter-temporal externalities and a short discussion in Sect. 5 concludes the paper.

2 Model

Let $\Gamma = (K, (A_i)_{i=1,2}, r, q)$ be a two-player zero-sum stochastic game, where K is the finite set of states, and A_i is the finite action set of Player i (assumed to be the same in each state). We denote the payoff function by $r : K \times A_1 \times A_2 \rightarrow \mathbb{R}_+$, so when actions $(a, b) \in A_1 \times A_2$ are chosen at state k , Player 2 pays to Player 1 the amount $r(k, a, b)$. The payoff is extended

¹ Since only one player controls the transitions, single-controller stochastic games are tractable and exhibit properties such as the existence of the value in stationary strategies [4], the order-field property [3, 10], and the existence of convenient algorithms for computation of the value ([11–13] and the algorithm presented in the Appendix.).

to mixed actions in the standard manner. We denote by $\|\Gamma\|_\infty = \max_{k,i,j} r(k, i, j)$ the maximum payoff in the game.

Time is discrete and denoted by $t = 0, 1, \dots$. The next state is determined according to the transition function $q : K \times A_1 \times A_2 \rightarrow \Delta(K)$. Hence, if actions (a, b) are played in state k , the probability that the state in the next stage is k' is $q(k'|k, a, b)$.

The payoff in this repeated game is the undiscounted average (lim inf).² More precisely, let s_t be the state at time $t = 0$ and a_t, b_t the pure actions played by the players (resp.) at time t . A history of length $t > 1$ is an element of $H_t = K \times (A_1 \times A_2 \times K)^t$ and includes the states and the actions played in the previous $t - 1$ stages, as well as the state at time t . At $t = 0$, the history is a singleton containing the initial state and $H_0 = K$. The strategy of Player i is a function $\sigma_i : \bigcup_{t=0}^\infty H_t \rightarrow \Delta(A_i)$. We define the payoff to be

$$\gamma(k_0, \sigma_1, \sigma_2) = \mathbb{E}_{\sigma_1, \sigma_2} \left(\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T r(k(t), a(t), b(t)) \right), \tag{1}$$

where $a(t), b(t)$ are chosen by σ_1, σ_2 according to the t -stage history, $k(t)$ is the state at time t , and the initial state is $k(0) = k_0$. We focus our discussion on two types of simpler strategies that do not depend on the entire history:

- A *stationary strategy* is a strategy that depends at time t on the state $k(t)$, but not on the rest of the history or on t itself. Formally, a stationary strategy is a strategy such that for each two histories of lengths t, τ (not necessarily $t = \tau$), $h = (k_0, a(0), b(0), \dots, k(t))$ and $\hat{h} = (\hat{k}_0, \hat{a}(0), \hat{b}(0), \dots, \hat{k}(\tau))$, if $k(t) = \hat{k}(\tau)$ then $\sigma(h) = \sigma(\hat{h})$. We denote by Σ_i the set of stationary strategies for Player i . Essentially, Σ_i is the set of functions from K to $\Delta(A_i)$.
- A *static strategy* is a strategy that dictates the same mixed action in each stage, independent of the state, t or the history (see [17]). Formally, a static strategy is a strategy such that for each two histories h, \hat{h} of any (possibly different) lengths, $\sigma(h) = \sigma(\hat{h})$. We denote by $\tilde{\Sigma}_i$ the set of static strategies for Player i . Essentially, $\tilde{\Sigma}_i$ is the set of constant functions so it can be identified with $\Delta(A_i)$.

We define the maximin value in stationary strategies to be

$$v(k_0) = \max_{\sigma \in \Sigma_1} \min_{\tau \in \Sigma_2} \gamma(k_0, \sigma, \tau), \tag{2}$$

where k_0 is the initial state of the game. The maximin value in static strategies $\tilde{v}(k_0)$ is similarly denoted by $\tilde{\Sigma}_1$ instead of Σ_1 . We assume that only Player 1 is limited to static strategies, so Player 2 can best respond with a stationary strategy (Example 1). This assumption is only relevant for Theorem 1, as in the case of repeated games with bonuses or switching costs, the best response of Player 2 to a static strategy is a static strategy. Let

$$\tilde{v}(k_0) = \max_{\sigma \in \tilde{\Sigma}_1} \min_{\tau \in \Sigma_2} \gamma(k_0, \sigma, \tau). \tag{3}$$

In both cases, this might not be *the value* of the stochastic game when we allow players to use history-dependent strategies. Still, for simplicity, we refer to them as the values in

² Filar [3] mentions (footnote 4) that following Bewley & Kohlberg there are six alternatives including lim inf (excluding discounting) to evaluate the averages, but all are equivalent when the optimal strategy is stationary, as it is in our model. We study the discounted and finite cases in a different paper [16] and show that for patient enough players the results converge to the lim inf-undiscounted payoff.

stationary and static strategies. Moreover, in the main part of the paper, we limit our attention to single-controller stochastic games (see below) where the value exists in stationary strategies [4], hence $v(k_0)$ is indeed the value of the game.

Example 1 A stochastic game where the best response of Player 2 to a static strategy of Player 1 is not static.

Consider the following two-state stochastic game, where Player 1 chooses rows, Player 2 chooses columns and the arrows represent the deterministic transitions after each pair of actions:

$$k_T : \begin{pmatrix} 0 \circlearrowleft & 0 \circlearrowleft \\ 1 \rightarrow & 0 \rightarrow \end{pmatrix} \quad k_B : \begin{pmatrix} 0 \leftarrow & 1 \leftarrow \\ 0 \circlearrowleft & 0 \circlearrowleft \end{pmatrix}$$

Suppose that Player 1 plays the static action $[x(T), 1 - x(B)]$ for $x \in (0, 1)$. The expected payoff of L in state k_T is $1 - x > 0$, and the expected payoff of R in state k_B is $x > 0$. Hence, the best response of Player 2 is to play R in state k_T and L in state k_B , with an overall payoff of 0. Any static strategy of Player 2 would result in a payoff strictly larger than 0.

Our main objective is to compare the optimal payoff that is obtainable using stationary strategies to the payoff that can be guaranteed using static strategies. Their difference is essentially the price of using history-independent strategies:

$$\Delta(k_0) = v(k_0) - \tilde{v}(k_0) \geq 0. \tag{4}$$

Without additional requirements regarding the stochastic game, limiting Player 1 to static strategies can be arbitrarily bad for him. For example, in the following game, the transitions are deterministic and determined solely by Player 1, but still he can guarantee only 0 in static strategies, as compared to 1 in stationary ones, since the transitions depend on the states.

Example 2 For a general stochastic game Γ , Δ can be arbitrarily bad, i.e., $\Delta(k_0) = \|\Gamma\|_\infty$.

Consider the following three-state stochastic game. Player 2 (chooses rows) has only one action and has no role in the game. Arrows represent the deterministic transitions, so k_R is an absorbing state and k_L is the initial state:

$$k_L : (0 \circlearrowleft, 1 \rightarrow) \quad k_M : (1 \leftarrow, 1 \rightarrow) \quad k_R : (0 \circlearrowleft, 0 \circlearrowleft)$$

It is clear that the optimal strategy for Player 1 is the stationary strategy that plays R at k_L and L at k_M , since it guarantees him the maximal payoff in this game: $v(k_L) = 1$. By contrast, a static strategy that plays L with probability x either (for $x < 1$) eventually reaches state k_R where the payoff is 0, or remains in k_L (for $x = 1$), and yields 0. It follows that any static strategy yields a payoff of $\tilde{v}(k_L) = 0$, and $\Delta(k_L) = 1$.

From now on, we consider only stochastic games where only Player 1 controls the states (*single-controller*), and the transitions are deterministic and *state-independent*. Since there are $|A_1|$ actions, this bounds the number of states that can be visited. If there are strictly less states than actions and several actions lead to the same state, we can artificially duplicate this state so that each action leads exactly to one state. Hence, without loss of generality we can assume $K = A_1$ and rename the states such that $q(a|\cdot, a, \cdot) = 1$, so states correspond to the previously played action of Player 1.

This class of games is denoted by \mathcal{G} , and for it $v(k_0)$ is indeed the value of the game, even when allowing players to use time-dependent strategies, as this is a subclass of single-controller stochastic games [4]. By construction, each state is reachable at the second stage and the payoff of the initial state is irrelevant in the long run, so hereafter we drop the initial state from the notation for v , \tilde{v} , and Δ .

3 Single-Controller Stochastic Games With Deterministic and State Independent Transitions

Our main result provides a bound on the difference between the payoff in static and stationary strategies for games in \mathcal{G} . The basic idea is to choose one “good” state where v can be obtained, and then play the static strategy that with probability 0.5 plays the action that leads to this state and with probability 0.5 plays the optimal action in this state. This ensures that half of the time the game is in the good state and in half of these occurrences, v is obtained. Thus, in quarter of the time, the value v is obtained, and therefore, Player 1 loses a maximum of $\frac{3}{4}v$. This is shown formally in the following Theorem, alongside an example that proves that the bound is indeed tight.

Theorem 1 Fix $\Gamma \in \mathcal{G}$. The loss incurred by Player 1 when using static strategies is at most $3/4$ of the value:

$$\Delta = v - \tilde{v} \leq \frac{3}{4}v$$

Remark 1 The previous result has an alternative interpretation in terms of ratio. The maximin value, \tilde{v} , that Player 1 can guarantee in static strategies is at least a quarter of the value v that he can guarantee in stationary ones.

Proof Let v be the value of the game in stationary strategies. There exists a state $k^* \in K$ and a mixed action $x^* \in \Delta(A_1)$ such that for all $y \in A_2$ we have $r(k^*, x^*, y) \geq v$ (otherwise, the payoff of Player 1 in each state cannot exceed v , as well as the value). Consider the static strategy σ_f of Player 1, which with probability 0.5 chooses the pure action that leads to the state k^* (named also k^*) and otherwise plays x^* . The payoff of this strategy against the stationary strategy $\tau(s)$ of Player 2 is

$$\begin{aligned} & \frac{1}{4}r(k^*, x^*, \tau(k^*)) + \frac{1}{4}r(k^*, k^*, \tau(k^*)) + \frac{1}{4} \sum_{k \in K} x^*(k)r(k, x^*, \tau(k)) \\ & + \frac{1}{4} \sum_{k \in K} x^*(k)r(k, k^*, \tau(k)) \geq \frac{1}{4}v + 0, \end{aligned}$$

where we bound the first term with v as stated before and the rest are bounded by the minimal payoff in the game, 0. Since this is one possible static strategy, its payoff is a lower bound on \tilde{v} and we obtain $\tilde{v} \geq \frac{1}{4}v$.

Example 3 The bound of Theorem 1 is tight.

Suppose $K = A_1 = \{0, \dots, 4\}$, $A_2 = \{1\}$ and the payoff function is $r(k, i, j) = \mathbb{1}_{i=k+1}$ where the summation is modulo 5, so $4 + 1 = 0$. The optimal stationary strategy is to choose in state i the action $i + 1$ (modulo 5) with $v = 1$. Let σ_f be a static strategy that chooses action i with probability x_i . Then, the payoff is $x_0x_1 + x_1x_2 + \dots + x_4x_0$ and it is straightforward to verify that the maximum of this function inside the simplex is $\frac{1}{4}$.

The above result gives a tight bound on the static value compared to the stationary value. For specific games, the bound can be tighter and the optimal static strategy can be different. In the next section, we consider stochastic games that arise from repeated games with bonuses/switching costs and use the additional structure to obtain a better bound. Another simple set of games where the bound is different are small games, i.e., games where $|A_1| \leq 4$. In this case, $\Delta \leq \frac{|A_1|-1}{|A_1|}v$, provided that Player 1 can guarantee at least v in each state (the optimal static strategy is to play the optimal action in each state with probability $\frac{v}{|A_1|}$). This is why in the above example she has 5 actions.

4 Repeated Games With Inter-Temporal Externalities

In this section, we present a special class of stochastic games, which are originated by adding inter-temporal externalities to one of the players in repeated normal form games. The resulting stochastic games are more structured, which enables us to improve the bound.

We provide a general model of repeated games with arbitrary structure of inter-temporal externalities generalizing both [15] and [17]. Two players play repeatedly the one-shot zero-sum game M . In addition to the M -related payoffs, at each stage Player 1 obtains a payoff from Player 2 depending on his previous action and his new action, so the stage payoff Player 2 pays to Player 1 is $M_{ij} + cB_{i'i}$, where $i(j)$ is the current action of Player 1 (2), i' is the action of Player 1 in the previous stage, B is a $|A_1| \times |A_1|$ matrix, and c is some positive constant. The matrix B can be of general structure; however, when $B_{ij} = 0$ except for $B_{ii} \geq 0$, the matrix B represents bonuses for repeating the previous action, as in [15], who studied the special case of B being a scalar matrix; when $B_{ij} \leq 0$ except for $B_{ii} = 0$, the matrix B represents switching costs as in [17].³ The constant c serves as a weight between the inter-temporal externalities and the one-shot game M , and allows us to study how the price of history-independent strategies changes as the relative weight between the two changes, as is done in [17]. We denote by $v(c)$ the value with stationary strategies, by $\tilde{v}(c)$ the value with static strategies, and $\Delta(c)$ their difference.

This formulation is equivalent to a single-controller stochastic game where the set of states is A_1 and the transitions are deterministic: after Player 1 played i at stage t , the state at stage $t + 1$ is i . The i, j element of the payoff matrix at state i' is $M_{ij} + cB_{i'i}$ and the payoffs are accumulated according to Eq. (1).

As for static strategies, in this model the payoff and the value can be written in a simple manner. If Player 1 uses a static strategy x and Player 2 uses the static strategy y , the payoff in matrix notation is $g(c)(x, y) = x^T M y + c x^T B x$. Hereafter, we drop the transpose symbol to simplify the notation.

We first show that these games are indeed more structured. Recall that \mathcal{G} is the set of single-controller stochastic games with deterministic and state-independent transitions and let \mathcal{G}_B be the set of stochastic games such that there exists two matrices M of size $|A_1| \times |A_2|$, and B of size $|A_1| \times |A_1|$ such that

$$\forall (i', i, j) \in K \times A_1 \times A_2 : r(i', i, j) = M_{ij} + B_{i'i} \text{ and } q(i|i', i, j) = 1. \quad (5)$$

The games in \mathcal{G}_B are also single-controller with deterministic and state-independent transitions, so $\mathcal{G}_B \subseteq \mathcal{G}$. Tsodikovich et al. [17] showed that for repeated games with switching costs (which are in \mathcal{G}_B), the best response to a static strategy is also a static strategy, whereas in Example 1 we showed that this is not true for all the games in \mathcal{G} . Thus, we conclude that $\mathcal{G}_B \neq \mathcal{G}$ and focus this section on \mathcal{G}_B , a subclass of games dealt with in Theorem 1.

The following characterization of \mathcal{G}_B can be used to see whether a game in \mathcal{G} is indeed in \mathcal{G}_B and how to construct the appropriate matrices M and B . The intuition behind this result is that there must be no interaction between the actions of the minimizing player and the states. Note that the ability to decompose a game in that manner is independent of c , as it is a multiplicative factor that can be absorbed into B , and we can assume that $c = 1$ for the rest of this discussion.

³ For clarity, when dealing with switching costs, we write explicitly Eq. (5) as $M - cS$, where S is the switching costs matrix with positive entries.

Proposition 1 *Let Γ be in \mathcal{G} , then Γ is in \mathcal{G}_B if and only if*

$$\forall i, i', i'' \in A_1, \forall j, j' \in A_2, r(i', i, j) - r(i'', i, j) = r(i', i, j') - r(i'', i, j'). \quad (6)$$

Proof The proof relies on the direct verification of Equation (6). Since the proof is technical, it is omitted.

We conclude that games in \mathcal{G}_B have more structure than games in \mathcal{G} , and with this additional structure the bound can be improved. The following Theorem shows that the bound from Theorem 1 is improved to half under the assumption that the game is in \mathcal{G}_B with bonuses, i.e., all entries of B are non-negative and symmetric⁴.

Theorem 2 *Fix $\Gamma \in \mathcal{G}_B$ such that B is symmetric with non-negative entries. The loss incurred by Player 1 when using static strategies is at most $1/2$ of the value:*

$$\Delta = v - \tilde{v} \leq \frac{1}{2}v$$

Remark 2 The previous result has an alternative interpretation in terms of ratio. The maximin value, \tilde{v} , that Player 1 can guarantee in static strategies is at least a half of the value v he can guarantee in stationary ones.

Proof We follow the same idea as in the proof of Theorem 1. Let v be the value of the game in stationary strategies. There exists an action $i^* \in A_1$ and a mixed action $x^* \in \Delta(A_1)$ such that for all $y \in A_2$ we have $r(i^*, x^*, y) \geq v$. Consider the static strategy τ_f of Player 1, which with probability 0.5 chooses the pure action i^* , and otherwise chooses an action according to mixed action x^* . Let y be a static best response of Player 2 to τ_f . The payoff when this strategy profile is played is

$$\begin{aligned} g(0.5x^* + 0.5i^*, y) &= (0.5x^* + 0.5i^*)My + (0.5x^* + 0.5i^*)B(0.5x^* + 0.5i^*) \\ &= \frac{1}{2}x^*My + \frac{1}{2}i^*My + \frac{1}{4}x^*Bx^* + \frac{1}{4}i^*Bx^* + \frac{1}{4}x^*Bi^* + \frac{1}{4}i^*Bi^* \end{aligned}$$

By assumption, $i^*Bi^* \geq 0$, $x^*Bx^* \geq 0$, and from symmetry $x^*Bi^* = i^*Bx^*$. Thus,

$$g(0.5x^* + 0.5i^*, y) \geq \frac{1}{2}x^*My + \frac{1}{2}i^*Bx^* \geq \frac{1}{2}v$$

where the last inequality follows from the fact that $r(i^*, x^*, y) = x^*My + i^*Bx^*$ is the expected payoff of playing x^* at some stage after playing i^* in the previous one.

Note that when some of the entries of B are negative (as in the case of switching costs), the stage payoffs can be negative, which prevents us from applying Theorems 1 and 2. However, these Theorems can be applied to games with bonuses as in [15] and its generalizations.

4.1 Optimality of a Static Strategy for Small and Large Values of c

In most models, bonuses due to repeating the previous actions or switching costs are either ignored or considered irrelevant, so comparing a case with no switching costs ($c = 0$) to one with small switching costs is important as a robustness check on this claim. Lipman & Wang [8] provide one such comparison in discussing the continuity of equilibrium payoffs and equilibrium strategies when switching costs converge to zero. Here, we provide another comparison: between the optimal stationary and static strategies when c is very small.

⁴ As pointed out in [16], this is not a limitation as it still generalizes previous studies [7, 8, 15] which assumed that non-diagonal elements of the matrix are identical, as in Sect. 4.1.1

For any M and B , clearly, $\Delta(0) = 0$, since for $c = 0$ there are no inter-temporal externalities and in a repeated zero-sum game there exists an optimal strategy which is static. Similarly, [17] showed in the switching costs model that $\Delta(c) = 0$ for all c whenever M is a 2×2 game. Obviously, this is not generally true. We start with two examples, to demonstrate how the two values behave as a function of c . Only the results are presented here, the calculations are in Appendix 1.

Example 4 A 3×3 repeated game with switching costs. Consider the following 3×3 version of matching pennies $M = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ with the switching costs matrix $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. For clarity, we do not normalize here the elements of M to be positive.

Without switching costs ($c = 0$), the optimal action for Player 1 is $\frac{1}{11}(6, 3, 2)$ and the value is $v = -\frac{6}{11}$. In the equivalent stochastic game, there are three states representing the situation after each of Player 1's actions. We denote by k_i the state where Player 1 played action $i \in \{T, M, B\}$ in the previous time period. In all states, the transitions are such that after action i was played by Player 1, the next state is k_i :

$$k_T : \begin{pmatrix} -1 & 0 & 0 \\ -c & -2 - c & -c \\ -c & -c & -3 - c \end{pmatrix} \quad k_M : \begin{pmatrix} -1 - c & -c & -c \\ 0 & -2 & 0 \\ -c & -c & -3 - c \end{pmatrix} \quad k_B : \begin{pmatrix} -1 - c & -c & -c \\ -c & -2 - c & -c \\ 0 & 0 & -3 \end{pmatrix}$$

The Value of the Game – $v(c)$

Case 1: $c < \underline{c} = \frac{22}{31}$: The optimal strategy for Player 1 in game M is still optimal for this case, and the value is $v(c) = -\frac{6}{11} - \frac{72}{121}c$.

Case 2: $\underline{c} \leq c \leq \bar{c} = \frac{121}{156}$: The switch from T to B is too costly, so action B is removed from state k_T and the optimal strategy becomes $[\frac{2}{3}(T), \frac{1}{3}(M)]$. The optimal strategies in the other states are the same as in the game without switching costs. The value is $v(c) = -\frac{156c+198}{319}$.

Case 3: $c \geq \bar{c}$: The switch from T to M is too costly, so k_T becomes an absorbing state with payoff -1 .

The Maximin Value in Static Strategies – $\tilde{v}(c)$

In this game, there is no intermediate strategy for Player 1. He plays the optimal strategy of the game M when switching costs are low enough, otherwise he plays the pure maximin strategy of M and never switches. A direct computation reveals that the cutoff c is $c_0 = \frac{55}{72}$.

Comparison: The maximal difference is at c_0 and is equal to $\Delta(c_0) = \frac{37}{2088}$.

In Example 4, small switching costs do not matter and the optimal stationary strategy is also a static one, but this is not the general case as the game $M = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^T$ with the switching costs matrix $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ demonstrates. Here, the value of M is 0.5 and can be achieved by any mixed action of Player 1 choosing the set of even rows with probability 0.5. The switching-cost matrix S is such that, based on the action of Player 1 in the previous time period, the only costless actions are the same action and the action above it. Due to this cyclical nature of S , Player 1 can achieve the value of game M with a stationary strategy, by playing the previous action w.p. 0.5 and the row above it w.p. 0.5. Clearly, any static non-pure strategy will contain a c component, and any pure static strategy will result in $\tilde{v}(c) = 0$. To conclude, $\tilde{v}(c) < v(c)$ for all $c > 0$.

Let \underline{c} be the maximal c such that for every $c \in [0, \underline{c}]$ the value is attained by a static strategy, i.e., $\tilde{v}(c) = v(c)$. If $\underline{c} > 0$, it immediately reveals the optimal strategy in the entire $[0, \underline{c}]$ region, both static and stationary. This optimal strategy in the game with switching

costs is repeatedly playing the optimal mixed action in the game M that provides the highest expected inter-temporal payoff.

Proposition 2 *Let \underline{c} be the maximal c such that $v(c) = \tilde{v}(c)$ for every $c \in [0, \underline{c}]$. If $\underline{c} > 0$, then for every $c \in [0, \underline{c}]$ the value is attained by the static strategy*

$$x^* = \arg \max_{x \in \mathcal{A}_1} xBx, \tag{7}$$

where \mathcal{A}_1 is the set of optimal strategies of Player 1 in the one-shot game M .

In addition, \underline{c} can be found in polynomial time.

Proof We show that x^* is an optimal stationary strategy for an interval of c s of the form $[0, \underline{c}]$ and explain how to find \underline{c} (which can also be $\underline{c} = 0$). The proof relies on the fact that the value in stationary strategies is piece-wise linear, and it is possible to find a strategy that is optimal for all c s in the entire segment (the proof from [17] for switching costs applies also for general externalities). Thus, there exists a strategy that is optimal for all $c \in [0, \underline{c}]$, and in particular for $c = 0$, i.e., a strategy in \mathcal{A}_1 . We show that the required strategy is x^* , the strategy with the lowest expected externalities among all the optimal strategies in M .

First, note that x^* is well defined: \mathcal{A}_1 is a compact set, xBx is continuous, so it has a maximum.

Second, we find the optimal strategy of Player 2. Let I_2 be the largest support (in terms of inclusion) of optimal actions of Player 2 in M and $y_0 \in I_2$ (arbitrarily chosen). The strategy y_0 is a best response against x^* for $c = 0$ but not necessarily for $c > 0$. We use the Average Cost Optimality Equation (ACOE, see Appendix 1) to find an optimal strategy.

In principle, ACOE is an algorithm used to solve single-controller stochastic games, yielding the value, optimal strategies and the continuation payoff from each state by solving a system of linear equations. Here, some of the above are already known, which significantly simplifies the computation. In particular, we assume that Player 1 uses x^* and Player 2 uses y_0 . The undiscounted payoff is therefore $\gamma(c) = v(0) + cx^*Bx^*$ and the only unknowns are the continuation payoffs in each state, which are linear functions of c .

After applying ACOE, plug in the continuation payoffs to each state and consider each state as a one-shot game to find the optimal actions of the players. For Player 1, the action x^* is still optimal as both the bonuses and the continuation payoffs rely only on his actions and they are the same in each column. When Player 1 plays x^* , he makes Player 2 indifferent among all the actions in I_2 and prefer them over all the other actions (as in the $c = 0$ case). Similarly, for Player 2, it is possible to solve a set of linear equations and inequalities to find if there exists a mixed action over I_2 (instead of y_0) that keeps Player 1 indifferent among all the actions for which $x_i^* > 0$ and still preferring them over the ones where $x_i^* = 0$. This results in a mixed action which depends on c in a linear manner and is optimal in the one-shot game corresponding to the state.

The stationary strategy of Player 2, which plays in each state the corresponding optimal mixed action, is optimal in the repeated game. It is left to verify that the mixed actions are indeed *distributions over the columns*, i.e., that all the columns are played with positive probabilities and that the sum of probabilities adds up to 1.⁵ For each pure action which is played with nonzero probability in each state this condition results in an upper bound on c found by solving a linear inequality. The minimum of all these upper bounds is \underline{c} .

In this proof, one solves one $|A_1| \times |A_1|$ linear equation (for the ACOE), then $|A_1|$ linear equations with $|A_2|$ variables and $|A_1|$ equations (for finding the mixed actions of Player 2).

⁵ This can be done while solving for the mixed action, but written here apart from that step for clarity, as this verification provides the value of \underline{c} .

Since the complexity of solving linear equations is polynomial, so is the entire algorithm. When $|A_2|$ is of the order of $|A_1|$, this is $O(|A_1|^4)$.

On the other extreme, the case of large c is interesting, as it represents a situation where the inter-temporal externalities are higher than any possible one-stage gain. For example, the profit of a firm that adjusts its price to dominate a market can be much smaller than the adjustment cost, provided that the other firms in the market respond quickly enough. Chakrabarti [2] studied this scenario for supergames and showed how it affects the set of possible equilibria and the optimal strategies, which must include only finitely many action changes.

In our model, without further assumptions, for large enough c , the optimal strategy can be completely dictated by the externalities. For example, if $B_{11} = 1$ and 0 otherwise, regardless of the game M , the optimal strategy for large c would be to repeat the action 1 and collect the bonus. Similarly, in a switching costs matrix similar to the one that appears after Example 4, the value of c has no effect on the value in stationary strategies but reduces the value in static ones to the pure maximin.

We therefore hereafter also assume that there are no free switches, i.e., all off-diagonal elements of B are dominated by the diagonal elements, so the payoff externalities for changing actions is strictly worse than for repeating pure actions. In the switching costs case, this means that the off-diagonal entries are strictly positive. In this case, increasing c increases the cost of playing mixed strategies and, at some point, it becomes optimal to play purely. The cutoff after which the optimal stationary strategy is pure is denoted by \bar{c} . The existence of \bar{c} was shown in [17] for switching costs and [15] for bonuses, and their proofs apply to the general case with no free switches.

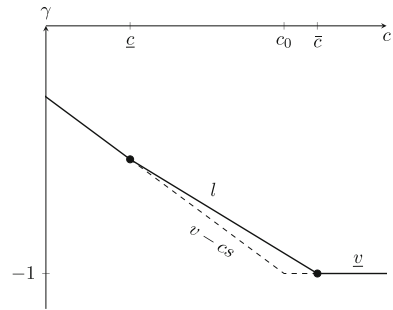
To avoid cumbersome notations and text, the rest of the section is formulated for games with switching costs. Similar results can be obtained for games with bonuses or a general inter-temporal externalities structure, with the proper adjustments. We denote the set of all games with switching costs and no free switches by \mathcal{G}_S , to emphasize the difference from games in \mathcal{G}_B with respect to the special structure of S compared to B .

Although we have no efficient algorithm to calculate \bar{c} , an upper bound for it can be found using a technique presented in [4]. They showed that to find the value of a single-controller stochastic game, one can consider a one-shot game where the pure actions are the pure stationary strategies of the stochastic game and the payoffs are the corresponding average payoffs. For example, the corresponding one-shot game to the one presented in Example 1 is a 4×4 game and the four pure actions of Player 1 are $\{TT, TB, BT, BB\}$, where the action ij chooses $i \in \{T, B\}$ in state k_T and $j \in \{T, B\}$ in state k_B . The main result of [4] states that the value of the one-shot game is the same as the value of the stochastic single-controller game, so we consider the one-shot game to bound \bar{c} :

A sufficient condition for the optimal stationary strategy to be pure is that one pure action of the one-shot game presented in [4] strictly dominates all others. Denote the i, j entry of this one-shot game matrix by $\mu_{ij} - \beta_i c$ (recall that only Player 1 controls the transition so the c -component is independent of j). Hence, every mixed entry in the matrix should suffice $\underline{v} \geq \mu_{ij} - \beta_i c$, where \underline{v} is the pure maximin. Assuming $\beta_i > 0$, we get $c \geq \frac{\mu_{ij} - \underline{v}}{\beta_i}$. When this condition is met for all i, j , we get $\tilde{v}(c) = \underline{v}$, hence $\bar{c} \leq \max_{i,j} \left(\frac{\mu_{ij} - \underline{v}}{\beta_i} \right)$.

As outside (\underline{c}, \bar{c}) the value functions are equal, the only region of interest is (\underline{c}, \bar{c}) . In this region, as a function of c , we can obtain a better bound based on the one-shot game and c itself. The following two theorems express the bounds on the loss $\Delta(c)$ due to playing by a static strategy.

Fig. 1 The value in optimal (stationary) strategies $v(c)$ (solid) and the maximin value in static strategies $\tilde{v}(c)$ (dashed) of the game from Example 4



Theorem 3 Fix $\Gamma \in \mathcal{G}_S$, and assume that all off-diagonal elements of S are strictly positive. The maximum loss that Player 1 incurs while using static strategies is smaller than

$$\Delta(c) = \begin{cases} (c - \underline{c}) \frac{\bar{c} - c_0}{\bar{c} - \underline{c}} s & \text{if } c \in (\underline{c}, c_0), \\ (\bar{c} - c) \frac{c_0 - \underline{c}}{\bar{c} - \underline{c}} s & \text{if } c \in (c_0, \bar{c}), \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $s = x^* S x^*$, and $c_0 = \frac{v - \underline{v}}{s}$ (\underline{c} and x^* are defined in Proposition 2).

Proof The proof of this theorem heavily relies on graphical intuition, which is provided in Fig. 1.

Define $s = x^* S x^*$ (x^* is defined in Eq. (7)). From convexity of $v(c)$ ($-v(c)$ is concave, [17]) lies below the line that connects $(\underline{c}, v - cs)$ and (\bar{c}, v) (denoted by l). The function $\tilde{v}(c)$ must be above the lines \underline{v} (corresponding to a pure static strategy) and $v - cs$ (corresponding to static strategy x^*). These lines meet at $c_0 = \frac{v - \underline{v}}{s}$. Hence, for $c < c_0$ the difference between the value in static and in stationary strategies is at most the difference between line $v - cs$ and l , and for $c > c_0$ it is the difference between \underline{v} and l . A direct computation reveals that these differences are as in Eq. (8).

In Theorem 3, we construct a bound by comparing an upper bound on $v(c)$ with a lower bound on $\tilde{v}(c)$. This lower bound is constructed by using two feasible and simple static strategies: the minimax pure action of game M and an optimal mixed action of game M , the one with the lowest switching cost. Any convex combination of these two static strategies is also a static strategy and can also be considered to improve the bound.

In particular, assume x^* is, for Player 1 in game M , the optimal mixed action with the lowest switching costs and \underline{x} is the pure maximin action of Player 1 (so $\underline{x} S \underline{x} = 0$). For every $\alpha \in [0, 1]$, $x_\alpha = \alpha x^* + (1 - \alpha) \underline{x}$ is also a feasible static strategy for Player 1 and its payoff bounds $\tilde{v}(c)$ from below.

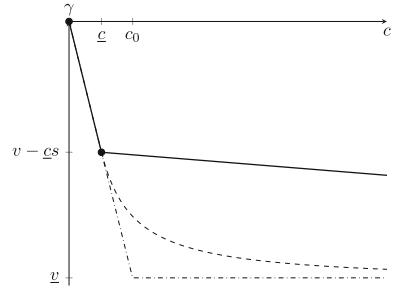
Proposition 3 Fix $\Gamma \in \mathcal{G}_S$, let \underline{x} be a pure maximin action in the one-shot game M , and denote by $\hat{s} = x^* S \underline{x} + \underline{x} S x^*$.⁶ If $s > \hat{s}$, then for every $c \in [c_1, c_2] \cap (\underline{c}, \bar{c})$, the bound from Theorem 3 can be improved by using

$$\Delta(c) = (v - \underline{v} + cs) \frac{\bar{c} - c}{\bar{c} - \underline{c}} - \frac{(v - \underline{v} + c\hat{s})^2}{4c(s - \hat{s})}, \tag{9}$$

where $c_1 = \frac{s}{2s - \hat{s}} c_0$ and $c_2 = \frac{\hat{s}}{s} c_0$.

⁶ Different bounds can be attained by using different pure maximin actions (if there are several). In that case, all of them can be constructed using the method presented, and the lowest chosen.

Fig. 2 The bound of Proposition 3, applied to Example 4. The solid line is $v(c)$, the dotted line is the bound (in this example, it is also $\tilde{v}(c)$) and the dot-dashed line is the bound from Theorem 3. For clarity, the graph was zoomed on the region near c_0 where the difference is the largest. All three meet at \bar{c}



Proof For every y , the payoff of the strategy pair (x_α, y) is

$$g(c)(x_\alpha, y) = x_\alpha My - cx_\alpha Sx_\alpha = \alpha x^* My + (1 - \alpha) \underline{x} My - c\alpha^2 x^* Sx^* - c\alpha(1 - \alpha) [x^* S\underline{x} + \underline{x} Sx^*],$$

where $c(1 - \alpha)^2 \underline{x} S\underline{x} = 0$ since \underline{x} is pure and there are no switches. Since x^* is optimal in M , it guarantees the value and $x^* My \geq v$. Similarly, \underline{x} guarantees the maximin value so $\underline{x} My \geq \underline{v}$. As previously, we let $s = x^* Sx^*$ and define $\hat{s} := x^* S\underline{x} + \underline{x} Sx^*$ (if S is symmetric, $\hat{s} = 2x^* S\underline{x}$). Therefore, $g(c)(x_\alpha, y)$ can be bounded by the following quadratic function of α :

$$g(c)(x_\alpha, y) \geq \alpha v + (1 - \alpha)\underline{v} - \alpha^2 cs - \alpha(1 - \alpha)c\hat{s} = -\alpha^2 c(s - \hat{s}) - \alpha(c\hat{s} - v + \underline{v}) + \underline{v} := f(\alpha).$$

Since $\tilde{v}(c) \geq g(c)(x_\alpha, y)$, it follows that $\tilde{v}(c) \geq \max_{\alpha \in [0, 1]} f(\alpha)$. Hence, by finding this maximum, we can obtain a tighter bound on $v(c) - \tilde{v}(c)$ relative to Theorem 3.

If f is convex, the maximum is obtained on one of the extreme points (depending on c), which takes us back to the case of Theorem 3. In fact, we depart from the case presented in Theorem 3 only if f is concave and the maximum is inside $(0, 1)$, i.e., only if $s > \hat{s}$ and $0 \leq \frac{v - \underline{v} - c\hat{s}}{2c(s - \hat{s})} \leq 1$.

Assume therefore that $s > \hat{s}$. The condition $0 \leq \frac{v - \underline{v} - c\hat{s}}{2c(s - \hat{s})} \leq 1$ can be rearranged as a condition on c . The “ $0 \leq$ ” part becomes $c \leq c_2 = \frac{v - \underline{v}}{\hat{s}}$ and the “ ≤ 1 ” becomes $c \geq c_1 = \frac{v - \underline{v}}{2s - \hat{s}}$.

Hence, for $c \in [c_1, c_2] \cap (\underline{c}, \bar{c})$, the bound from Theorem 3 can be improved by using

$$f\left(\frac{v - \underline{v} - c\hat{s}}{2c(s - \hat{s})}\right) = \underline{v} - \frac{(v - \underline{v} - c\hat{s})^2}{4c(s - \hat{s})}$$

instead of $\max\{\underline{v}, v - cs\}$. Thus, the bound for this domain is

$$\Delta(c) = (v - \underline{v} - cs) \frac{\bar{c} - c}{\bar{c} - \underline{c}} - \frac{(v - \underline{v} - c\hat{s})^2}{4c(s - \hat{s})}.$$

An illustration of this bound can be found in Fig. 2. Interestingly, in this case, the optimal static strategy is actually a mixture of x^* and \underline{x} , so the estimate done in Proposition 3 of $\tilde{v}(c)$ is accurate. We note that in the special case of action-independent switching costs (Sect. 4.1.1) the conditions for Proposition 3 never hold (Observation 1). Thus, this bound is applicable only in the case that there is enough variance in the switching costs matrix.

4.1.1 Action-Independent Inter-Temporal Externalities

In this section, we consider the special case of inter-temporal externalities which are independent of the action, i.e., $B_{ij} = 0$ and $B_{ii} = 1$ (for switching costs, $s_{ij} = 1$ for all $i \neq j$. Note that here the two models are strategically equivalent.). This case naturally arises when the switching costs stem from the act of “switching” itself and do not depend on the actions being switched to or from, and it is the baseline in most switching costs literature [7, 8, 15]. In particular, in the case of static strategies, this assumption implies that $\tilde{v}(c)$ is a piece-wise linear function and that the optimal static strategy remains optimal for small changes of c (more precisely: for changes of c that keep it in the same segment of linearity of $\tilde{v}(c)$, see Theorem 2 in [17]). Moreover, Schoenmakers et al. [15], provided a way to compute the optimal static strategy whenever M is a square matrix. The explicit expression of the switching costs function allows us to derive a bound on the optimality of the value in static strategies for this case.

Corollary 1 *When switching costs are action-independent ($s_{ij} = 1$ for $i \neq j$ and 0 otherwise), the maximum loss that Player 1 can incur while using static strategies is*

$$\Delta(c) \leq \frac{v(c)}{2} + c \left(1 - \frac{1}{n}\right), \tag{10}$$

where $n = |A_1|$, the number of actions Player 1 has.

Proof Since all the switching costs are identical, $\max_{x \in \Delta(A_1)} xSx = \max_{x \in \Delta(A_1)} 1 - \|x\|^2$. Inside the simplex, the maximum is achieved at $x = (\frac{1}{n}, \dots, \frac{1}{n})$ and is equal to $1 - \frac{1}{n}$. Plugging these into the proof of Theorem 2, the bound becomes $\tilde{v}(c) \geq \frac{v(c)}{2} - c \left(1 - \frac{1}{n}\right)$.

Interestingly, this bound cannot be improved by harnessing Proposition 3 to this case. The reason is that the condition $s > \hat{s}$ can never hold in this kind of switching-costs matrix, and requires inequality of some kind in the switching costs.

Observation 1 *When switching costs are action-independent, the condition $s > \hat{s}$ from Proposition 3 never holds.*

Proof By definitions, $s = 1 - \|x^*\|^2$, and $\hat{s} = 2(1 - \underline{x} \cdot x^*)$. The condition $s > \hat{s}$ from Proposition 3 becomes $1 - \|x^*\|^2 > 2(1 - \underline{x} \cdot x^*)$ or equivalently $(2\underline{x} - x^*) \cdot x^* > 1$. However, the maximum of the function $(2\underline{x} - z) \cdot z$ (over all \mathbb{R}^n) is at $z = \underline{x}$, and is equal to $\|\underline{x}\|^2 = 1$. Thus, $(2\underline{x} - x^*) \cdot x^* \leq 1$ and the desired inequality never holds.

5 Discussion

We study stochastic games where the players are restricted to using history-independent strategies, namely *static strategies*. We obtain a bound on the price of being static in the general case and then, in a series of theorems, improve it to specific cases with additional structures. This additional structure improves the value that can be obtained by static strategies, especially when the switching costs are relatively low. For example, when considering games resulting from repeated games with switching costs, static strategies can ensure (almost) half of the value (Theorem 2), a bound that can be improved significantly with additional knowledge regarding the game (Theorem 3).

Author Contributions All authors contributed equally.

Funding Open access funding provided by Bar-Ilan University. YT acknowledges the support of the Israel Science Foundation grants #2566/20, #1626/18 & #448/22. XV is a member of GNAMPA-INdAM. This research project received partial support from the Italian MIUR PRIN 2017 Project ALGADIMAR “Algorithms, Games, and Digital Markets” and the MIUR PRIN 2022 project “Learning in Markets and Society” (project 2022EKNE5K).

Data Availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval. Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Average Cost Optimality Equation (ACOE)

In this Appendix, we show how the value of the game and the optimal stationary strategies can be computed using the Average Cost Optimality Equation. We achieve this by solving Example 4 in detail. More about this method and proofs can be found in [1], the references within, and similar papers.

In general terms, the algorithm consists of the following steps:

1. “Guess” a stationary strategy for Player 1. Based on the ideas of Theorem 1 in [17], this stationary strategy should in each state play a mixed action that keeps Player 2 indifferent among several of his actions and makes him prefer them over the rest in the one-shot game M . This well-defines a finite set of actions to consider and ensures that the algorithm eventually terminates.
2. Player 2 does not control transitions, so once the strategy of Player 1 is determined, he can best respond by myopically best responding in each state. This can be achieved even by pure actions. Fix such a best response.
3. Let q_{ij} be the probability of Player 1 playing action j in state k_i , let r_i be the value of one-shot game M when the players use their actions chosen for state k_i , let v_i be the continuation payoff of the repeated game starting from state k_i and let γ be the value of the game (Eq. (1)). The ACOE connects these variables in the following manner:

$$\gamma + v_i = r_i - c \sum_j q_{ij} s_{ij} + \sum_j q_{ij} v_j, \quad (11)$$

where s_{ij} is an element in the matrix S .

4. The above equation actually represents n linear equations (one for each state) with $n + 1$ variables: v_1, \dots, v_n and γ . W.l.o.g. set $v_1 = 0$ and solve the remaining equations to obtain the value of the game γ and the continuation payoffs v_i .
5. Find the real mixed actions of Player 2. To do so, treat each state as a one-shot game by writing the switching costs and the continuation payoffs in the payoff matrix. Check

that Player 2 has a best response to Player 1’s action in this game that keeps Player 1 indifferent among the actions he plays and makes him prefer them over all other actions.

- If it is possible to find such a strategy for Player 2, this pair of strategies is optimal and γ is the value of the game. Otherwise, restart the algorithm with another guess.

A convenient way to see that the algorithm is finite is to assume the initial guess is a stationary strategy that mixes all the actions of Player 1 (in each state), and if it fails, use another that does not use one pure action, then two pure actions and so on. Eventually, the process terminates as there are only finitely many possible pure actions to stop considering.

We now apply this algorithm to the zero-sum game $M = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ and the switching-costs matrix $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Without switching costs ($c = 0$), the optimal strategy for Player 1 is $\frac{1}{11}(6, 3, 2)$ and the value is $v = -\frac{6}{11}$.

Suppose that Player 1 plays $\frac{1}{11}(6, 3, 2)$ in all states. This makes Player 2 indifferent among all his actions, and we can assume that the best response is simply playing the first column. This results in $r_i = -\frac{6}{11}$ and $(q_{iL}, q_{iM}, q_{iR}) = \frac{1}{11}(6, 3, 2)$ for all i . The three ACOE equations are

$$\begin{aligned} \gamma + v_L &= -\frac{6}{11} - c\frac{5}{11} + \frac{6}{11}v_L + \frac{3}{11}v_M + \frac{2}{11}v_R, \\ \gamma + v_M &= -\frac{6}{11} - c\frac{8}{11} + \frac{6}{11}v_L + \frac{3}{11}v_M + \frac{2}{11}v_R, \\ \gamma + v_R &= -\frac{6}{11} - c\frac{9}{11} + \frac{6}{11}v_L + \frac{3}{11}v_M + \frac{2}{11}v_R. \end{aligned}$$

Setting $v_L = 0$ and solving these equations leads to $v_M = -\frac{3c}{11}$, $v_R = -\frac{4c}{11}$ and $\gamma = -\frac{6}{11} - \frac{72}{121}c$. Our final step is to add these continuation payoffs to the payoff matrix of each state and find optimal strategies for Player 2. The resulting three one-shot games are as follows:

$$\begin{aligned} k_T &: \begin{pmatrix} -1 & 0 & 0 \\ -\frac{14c}{11} & -2 - \frac{14c}{11} & -\frac{14c}{11} \\ -\frac{15c}{11} & -\frac{15c}{11} & -3 - \frac{15c}{11} \end{pmatrix} & k_M &: \begin{pmatrix} -1 - c & -c & -c \\ -\frac{3c}{11} & -2 - \frac{3c}{11} & -\frac{3c}{11} \\ -\frac{15c}{11} & -\frac{15c}{11} & -3 - \frac{15c}{11} \end{pmatrix} \\ k_B &: \begin{pmatrix} -1 - c & -c & -c \\ -\frac{14c}{11} & -2 - \frac{14c}{11} & -\frac{14c}{11} \\ -\frac{4c}{11} & -\frac{4c}{11} & -3 - \frac{4c}{11} \end{pmatrix} \end{aligned}$$

In the game k_T , solving the standard equation of indifference leads to a mixed strategy in which Player 2 plays his actions (from left to right) with probabilities $\frac{1}{121}(72c + 66, 33 - 41c, 22 - 31c)$ (and similar solutions for the other states). This mixed action is indeed a distribution for $c \leq \frac{22}{31}$ and for these values of c we have found the value $v(c) = -\frac{6}{11} - \frac{72c}{121}$ and the optimal strategies.

For $c > \frac{22}{31}$, action B in state k_T becomes dominated and Player 2 cannot keep Player 1 indifferent between it and the other actions. We therefore assume Player 1 never plays action B in state k_T . As a result, Player 2 will never play his action R in k_T and the new guess for a mixed action for Player 1 in state k_T is $\frac{1}{3}(2, 1, 0)$. In the rest of the states, action B is not dominated so Player 1 continues to play mixed action $\frac{1}{11}(6, 3, 2)$ in them.

Repeating the same process with this strategy (which is not static this time) leads to the next critical value of $c = \frac{121}{156}$. Above it, action M in state k_T (including, again, continuation payoffs and switching costs) is dominated by T and so k_T becomes absorbing with $v(c) = -1$.

References

1. Arapostathis A, Borkar VS, Fernández-Gaucherand E, Ghosh MK, Marcus SI (1993) Discrete-time controlled Markov processes with average cost criterion: a survey. *SIAM J Control Optim* 31(2):282–344
2. Chakrabarti SK (1990) Characterizations of the equilibrium payoffs of inertia supergames. *J Econ Theor* 51(1):171–183
3. Filar JA (1981) Ordered field property for stochastic games when the player who controls transitions changes from state to state. *J Optim Theory Appl* 34(4):503–515
4. Filar JA, Raghavan TES (1984) A matrix game solution of the single-controller stochastic game. *Math Oper Res* 9(3):356–362
5. Filar JA, Schultz TA (1986) The traveling inspector model. *Oper Res Spektr* 8(1):33–36
6. Filar J (1985) Player aggregation in the traveling inspector model. *IEEE Trans Autom Control* 30(8):723–729
7. Lipman BL, Wang R (2000) Switching costs in frequently repeated games. *J Econ Theor* 93(2):149–190
8. Lipman BL, Wang R (2009) Switching costs in infinitely repeated games. *Games Econ Behav* 66(1):292–314
9. 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*. <https://doi.org/10.1007/s10589-021-00282-7>
10. Parthasarathy T, Raghavan TES (1981) An orderfield property for stochastic games when one player controls transition probabilities. *J Optim Theory Appl* 33(3):375–392
11. Raghavan TES (2003) Finite-step algorithms for single-controller and perfect information stochastic games. In *Stochastic games and applications*, pp 227–251. Springer
12. Raghavan TES, Filar JA (1991) Algorithms for stochastic games—a survey. *Z Oper Res* 35(6):437–472
13. Raghavan TES, Syed Z (2002) Computing stationary nash equilibria of undiscounted single-controller stochastic games. *Math Oper Res* 27(2):384–400
14. Rass S, König S, Schauer S (2017) Defending against advanced persistent threats using game-theory. *PLoS ONE* 12(1):e0168675
15. Schoenmakers G, Flesch J, Thuijsman F, Vrieze OJ (2008) Repeated games with bonuses. *J Optim Theory Appl* 136(3):459–473
16. Tsodikovich Y, Venel X, Zseleva A (2024) Folk theorems in repeated games with switching costs
17. Tsodikovich Y, Venel X, Zseleva A, Tsodikovich Y, Venel X, Zseleva A (2023) The regularity of the value function of repeated games with switching costs. *Math Oper Res* 48(4):1899–1905
18. Wachter J, Rass S, König S (2018) Security from the adversary's inertia-controlling convergence speed when playing mixed strategy equilibria. *Games* 9(3):59

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.