



# The Generalized Stackelberg Equilibrium of the Two-Person Stopping Game

Marek Skarupski<sup>1</sup> · Krzysztof J. Szajowski<sup>1</sup>

Received: 28 April 2022 / Revised: 14 December 2022 / Accepted: 3 February 2023 /

Published online: 10 May 2023

© The Author(s) 2023

## Abstract

In modeling the bilateral selection of states of the process, Dynkin (Dokl Akad Nauk USSR 185:241–288, 1969) proposed a two-person game in which players use stopping moments as strategies. The purpose of this work is to present a model of the game in which the players have different information about the process itself, as well as various laws to stop the process and accept its state. The game model uses the stochastic process apparatus, in particular, the ability to create different filters for the same process. The sets of stopping moments based on different filters are not identical, which allows us to model different sets of strategies for players. We show that the follower, by observing the behavior of a rational leader, can recover information that is lost due to the lack of complete observation of the state of the process. In the competition of two opponents for the maximum of the i.i.d. sequence, one of whom has access to full information and the other only knows their relative ranks, we found the generalized Stackelberg equilibrium. If the priority of a player observing the relative ranks is less than 50%, then that player modifies his strategy based on the behavior of the second player. For a player with full information, information about the behavior of the player observing the relative ranks is useless.

**Keywords** No-zero-sum game · Optional stopping · Equilibrium · Filtration

**Mathematics Subject Classification** 90D15 · 60G40 · 93C30

---

The research has been supported by Wrocław University of Science and Technology, Faculty of Pure and Applied Mathematics (No. 8211204601 MPK: 9130740000).

---

✉ Marek Skarupski  
Marek.Skarupski@pwr.edu.pl  
Krzysztof J. Szajowski  
Krzysztof.Szajowski@pwr.edu.pl

<sup>1</sup> Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, 50-370 Wrocław, Dolonoslaskie, Poland

## 1 Introduction

The very well-known secretary problem also has many modifications. Ferguson in [1] has made a review of the concepts of the best-choice problem (BCP) going back to the age of Kepler and the paper [2] by Cayley. Presman and Sonin in [3] considered so called no-information problem in which the appearing objects come from the rank distribution, that is, the objects are observable, the decision maker can rank them, and all permutations of the appearing objects are equally possible. The review paper [4] by Gilbert and Mosteller presented various models, also such that the exact value of the object is observable and the distribution of the object is known. (It is assumed to be a uniform distribution in the interval  $[0, 1]$ .) Both ideas can be described as the optimal stopping of a Markov chain. In both, there is only one decision maker and there is no competition concept. The game theory approach to the secretary problem was introduced by Dynkin and presented in [5]. The problem of choosing the best object is later used to show the role of information that decision makers have. As in our work [6], we are dealing with a bilateral decision problem related to the observation of the Markov process by decision makers. The information provided to the players is based on the aggregation of the observation data. Acceptable strategies are moment of hold related to the available information. The payouts are the result of the selected state at the time that the decision maker stopped observing. As in the cited work, the decision makers are not identical (symmetric). They differ in access to information and have different rights to access the observed state, similar to the Stackelberg's model (cf. [7]). The considered model is an extension of Dynkin's game, but is also closely related to the models presented in the works [8–10], or [11]. Further, examples can be found in Mazalov's book [12]. Another way to agree on acceptable behavior in stopping games has been proposed in the papers [13–16] (also cf. [17]).

### 1.1 Business Motivation

Consider a two companies A and B. Both of them are interested in buying a bundle of certain products on a commodity exchange. Company A is a large corporation and knows the actual value of the product on the market. In addition, it knows the previous values of the objects and can compare them. The problem of company B is that it does not have information about the actual value of the good. However, the owner of company B can compare the actual position of the good in the market with the previous observations. Both players want to choose the very best object overall without the possibility of recall. The number of objects is fixed and finite. A very good example can be described from the reliability position. Consider two buyers of the same item. Both want to buy the most reliable object. Buyer A has the ability to know the values of the reliability function derived by experts and quality controllers. The player B has no such contact and intelligence, so he must rely on his basic knowledge and the knowledge of the previous observation, i.e., he can judge whether the object is better or worse than the previous one. We can say that the buyers of the objects are two types: the first is a business, and the decision problem is preceded by unilateral consideration. The form of the optimal strategy in the decision problem

is the inspiration for the mean-value formulation. Threshold strategies are crucial tools for optimal stopping problems. The simplest case related to the observation of a sequence of random variables can be found, e.g., in [18] or [19]. The bilateral extension of these models can be found in [20]. Two players, **I** and **II**, observe sequentially a known finite number (or a number having a geometric distribution) of independent and identically distributed random variables. They must choose the largest. Variables cannot be perfectly observed. When a random variable is sampled, the sampler is informed only whether it is greater or less than some level that he has specified. Each player can choose at most one observation. After the sampling, the players decide on acceptance or rejection of the observation. If both accept the same observation, Player **I** has priority. The class of adequate strategies and a gain function are constructed. In the finite case, the game has a solution in pure strategies. In the case of a geometric distribution, Player **I** has a pure equilibrium strategy, and Player **II** has either a pure equilibrium strategy or a mixture of two pure strategies. The game is symmetric, as the players are watching the same string to the same extent. Increasing opposing interests is possible by completely different preferences of the players. Evaluation of the same object by two decision makers can mean that players observe the different coordinates of the vector and formulate their expectations for their realization. When players' aim is to achieve a minimum level of the observed rate, then the problem can be reduced to a game in which strategies are setting of just levels. Discussion of such issues can be found in the Sakaguchi's works (e.g., [21]). However, in those tasks though, the information players are incomplete, lacking clear asymmetry players. The pay-offs of the players are function of the thresholds, and the perfect comparison of the observed variable with these defined levels is guaranteed. Asymmetric tools in measure of the observed r.v. are presented in [22]. However, for private random variables, these players with asymmetric tools applied to their same sequence are the subject of consideration in the paper.

## 1.2 Mathematical Formulation of the Problem

In fine-tuning the mathematical model, we will use the methods of optimal stopping of stochastic processes presented in the monograph by Chow et al. [23] (for Markov sequences in the monograph by Dynkin and Yushkievich in [24, 25]), and Dynkin's game models with optimal stopping [5] of such sequences, similar to what is done in the works [10, 26].

Let  $(\Omega, \mathcal{F}, \mathbf{P})$  be rich enough probability space to define the random sequence  $\{X_n\}_{n=0}^N$ ,  $X : \Omega \rightarrow \mathbb{R} \subset \mathfrak{R}$ ,  $N \in \mathbb{N} \cup \{\infty\}$ . In general, one can define the filtration  $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$  and the set of stopping times  $\mathfrak{S}$  with respect to the filtration. There are two observers (and, at the same time, decision makers) of the basic sequence defined by the mappings  $\{\varphi_n^i\}$ ,  $i = 1, 2$ , where  $\varphi_n^i : \mathfrak{R}^n \rightarrow \mathfrak{R}$ , having his objectives defined by the payoff functions  $f^i : \mathbb{R} \rightarrow \mathfrak{R}$ . In other words, the player **I** at the moment  $n$  observes  $\xi_n^i = \varphi_n^i(X_1, \dots, X_n)$ . Let us denote  $\mathfrak{S}^i$ ,  $i = 1, 2$ , the sets of stopping times with respect to the filtration  $\mathcal{F}_n^i = \sigma\{\xi_1^i, \dots, \xi_n^i\}$ . The strategies of the players are stopping times  $\tau \in \mathfrak{S}^i$ . Each player, on the basis of the observations available to him, is tasked with choosing the moment of accepting the state of the

process based on the previous observations to maximize the expected payment.

$$\hat{v}^i = \sup_{\tau^i \in \mathfrak{S}^i} \mathbb{E} f^i(X_{\tau^i}). \tag{1}$$

It can be used to reduce the initial problem to the task of optimal stopping of conditional expected values relative to its filtration (v. [25]). Let us calculate for every  $n \in \mathbb{N}$

$$\hat{f}^i(\vec{\xi}_n^i) = \mathbb{E}[f^i(X_n^i) \mid \mathcal{F}_n^i],$$

where  $\vec{\xi}_n^i = (\xi_1^i, \dots, \xi_n^i)$ . We have

$$\hat{v}^i = \sup_{\tau^i \in \mathfrak{S}^i} \mathbb{E} \hat{f}^i(\vec{\xi}_{\tau^i}^i).$$

Let us assume that the observation processes  $\{\xi_n^i\}_{n=0}^N, i = 1, 2$ , belong to Markov processes. In this case, the solution of the problem (1) can be obtained using the procedure described in [25, Ch.3] which is based on the Bellman-Jacobi equation. Denote  $\mathfrak{S}_n = \{\tau \in \mathfrak{S} : \tau \geq n\}$ . When we have two decision makers hunting for a convenient state of the process, they have the right to declare stopping at most twice. The second one is when the first hired state is assigned to the opponent. The natural set of strategies are  $\mathfrak{U}^i = \{(\tau^i, \{\sigma_n^i\}_{n=0}^N) : \tau^i \in \mathfrak{S}, \sigma_n^i \in \mathfrak{S}_n^i\}$ . The pay-off in the competitive case is defined in various ways. Following the discussion of the paper [10] for given  $\rho^i \in \mathfrak{U}^i, i = 1, 2$ ,

$$\begin{aligned} \mathbf{K}_1(\rho^1, \rho^2) = & \mathbb{E} \left[ \mathbb{I}_{\{\tau^1 < \tau^2\}} \left( \hat{f}^1(\xi_{\tau^1}^1) - \hat{v}^2(\tau^1, \xi_{\sigma_{\tau^1}^2}^2) \right) \right. \\ & + \mathbb{I}_{\{\tau^1 = \tau^2\}} \left[ p \left( \hat{f}^1(\xi_{\tau^1}^1) - \hat{v}^2(\tau^1, \xi_{\sigma_{\tau^1}^2}^2) \right) \right. \\ & \left. \left. + (1 - p) \left( \hat{v}^1(\tau^2, \xi_{\sigma_{\tau^2}^1}^1) - \hat{f}^2(\xi_{\tau^2}^2) \right) \right] \right. \\ & \left. + \mathbb{I}_{\{\tau^1 > \tau^2\}} \left( \hat{v}^1(\tau^2, \xi_{\sigma_{\tau^2}^1}^1) - \hat{f}^2(\xi_{\tau^2}^2) \right) \right], \tag{2} \end{aligned}$$

where  $\mathbb{I}_A$  is the characteristic function of  $A$  and

$$\hat{v}^i(n, \xi_n^i) = \sup_{\tau^i \in \mathfrak{S}_n^i} \mathbb{E} \hat{f}^i(\vec{\xi}_{\tau^i}^i),$$

and  $0 \leq p \leq 1$  is the priority parameter, i.e., the probability that the state will be assigned to Player I. The pair of strategies  $(\rho^{1*}, \rho^{2*})$  is the solution to the problem if for every  $\rho^i \in \mathfrak{U}^i$ ,

$$K_1(\rho^{1*}, \rho^{2*}) \geq K_1(\rho^1, \rho^{2*}) \text{ and } K_1(\rho^{1*}, \rho^{2*}) \geq K_1(\rho^{1*}, \rho^2).$$

In practice, it is difficult to construct the solution and calculate the value of the problem in such general form. However, for some natural cases, each player can estimate his final reward by calculating his potential reward (award) based on his knowledge (filtration). The idea of these simplifications is presented in the next sections (Fig. 1).

## 2 Formulation of the Game Related to BCP

### 2.1 The Description of the Model

Consider a game in which two players want to choose the best object overall. They observe  $N$  objects sequentially. They get a profit only if the player chooses the best object and the rival will not do it.<sup>1</sup> In other case, he does not get the award. At each moment players analyze the state in sequential order and declare his wish. The second player claims his decision first. If both players want to stop on the current object, the nature chooses beneficiary by a lottery. Suppose that

- 1 The player **I** have no information, i.e., at any time prior to the decision, he observes only the relative ranks of the current objects and the behavior of player **II**.
- 2 The player **II** has full information, i.e., he observes sequentially  $X_1, \dots, X_N$  i.i.d., sees its value, and also can calculate the rank of the current object, and he presents the decision as soon as the observation is taken that player **I** knows it before making his own.
- 3 A player who accomplishes his goal (chooses the moment when the process reaches its global maximum) gets a payout of 1. If his pick is unsuccessful, he incurs a  $-1$  penalty. Otherwise, he ends the game with a payout of 0.

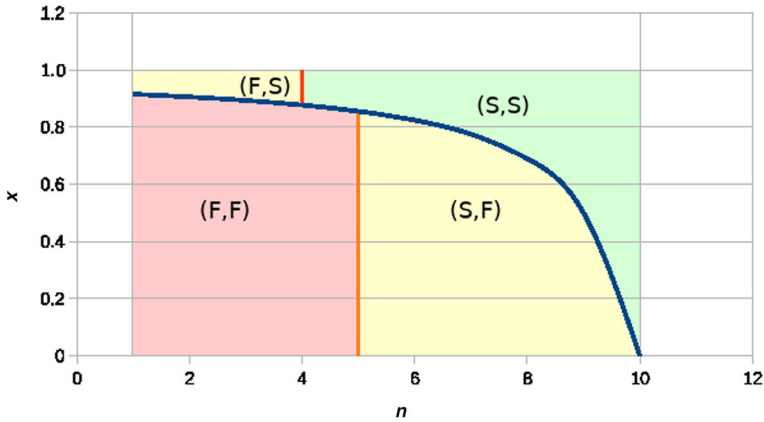
To be more specific, let us denote by  $Y_n$  the relative rank of the  $n$ -th observation

$$Y_n = \#\{1 \leq i \leq n : X_i \leq X_n\}. \tag{3}$$

Player **II** filtration is  $\mathcal{F}_n^{(1)} = \sigma(X_1, \dots, X_n)$  and for Player **I** it is  $\mathcal{F}_n^{(2)} = \sigma(Y_1, \dots, Y_n)$ . Note that  $\mathcal{F}_n^{(2)} \subset \mathcal{F}_n^{(1)}$  for every  $n$ . Denote by  $\mathcal{T}_1$  a set of all stopping times with respect to family  $\{\mathcal{F}_n\}_{n=1}^N$ . Let  $\mathcal{T}_1^0$  denote a set of all stopping times  $\tau \in \mathcal{T}_1$  such that  $X_n = \max\{X_1, \dots, X_n\}$  on  $\{\tau = n\}$ ,  $n = 1, \dots, N$ , and  $\mathcal{T}_{1,n} = \{\tau \in \mathcal{T}_1 : \tau \geq n\}$ . Define the moments where the greatest observations appear, that is,  $\tau_1 = 1$ ,  $\tau_k = \inf\{n : \tau_{k-1} \leq n \leq N, X_n = \max\{X_1, \dots, X_n\}\}$  for  $k = 1, \dots, N$ . We observe the sequence  $\tau_1, \tau_2, \dots \in \mathcal{T}_1^0$ . Now let us consider the following chain,

$$Z_k = (\tau_k, X_{\tau_k}) \text{ on } \{\tau_k < N + 1\}, Z_k = (\tau_{N+1}, \partial),$$

<sup>1</sup> More sophisticated approaches, as, e.g., player’s premia for finding the better one than opponent, will be to discuss in conclusions.



**Fig. 1** Boundaries of the strategies for  $N = 10, p = 0.25$ . The shift for Player I is clearly visible. In this case  $n^* = 4$  but  $\tilde{n} = 5$

where  $\partial$  is a special absorbing state. It is easy to see that  $\{Z_k\}_{k=1}^{N+1}$  is a Markov chain with transition probabilities (cf. [27])

$$p((n, x), (m, B)) = x^{m-n-1} \int_B dy, \text{ for } m > n, x \in (0, 1],$$

and, 0 otherwise, with  $B \subseteq (x, 1]$ . It means that the density function

$$p((n, x), (m, dy)) := p((n, x), (m, (y, y + dy))) = x^{m-n-1} dy, \tag{4}$$

for  $m > n, x, y \in (0, 1], x \leq y$ .

The reward for Player II for stopping at the  $n$ th object of the value  $X_n = x$  is the following

$$s_{2,n}(x) = x^{N-n}, \tag{5}$$

and for continuing observation and stopping on the next local maximum, taking into account (4), it is given by (cf. [4, 27])

$$c_{2,n}(x) = \sum_{k=n+1}^N \int_x^1 s_{2,k}(y) p((n, x), (k, dy)) = \sum_{k=n+1}^N \frac{x^{k-n-1} (1 - x^{N-n})}{N - k + 1}.$$

Since the sequence of local maxima is increasing, then  $c_{2,n}(x) \leq s_{2,n}(x)$  for  $x \geq x_n$ , where  $x_n$  is the only solution to the equation  $c_{2,n}(x) = s_{2,n}(x) \in (0, 1]$ . An explicit form of the equation is

$$\sum_{j=1}^{N-n} \frac{x^{-j} - 1}{j} = 1.$$

Based on the above denotations and theory of optimal stopping, we have

$$u_{2,n}(x) = \sup_{\tau \in \mathcal{T}_{1,n}^0} \mathbf{E}_{(n,x)} s_{2,\tau}(X_\tau) = \max\{s_{2,n}(x), \mathbf{T}u_{2,n}(x)\}. \tag{6}$$

Similarly, for the player I, consider a sequence of indicators  $\{I_n\}_{n=1}^N$ , where  $I_k = \mathbb{I}_{\{Y_k=1\}}$ . Let us denote by  $\mathcal{G}_n = \sigma(I_1, \dots, I_n)$  sequence of sigma fields generated by indicators, and let  $\mathcal{T}_2$  be the set of all stopping moments  $\tau$  with respect to  $\sigma$ -fields  $\mathcal{G}_n$ ,  $n = 1, \dots, N$ . Define a process  $\xi_t$  in the following way,

$$\xi_t = \inf\{n \geq \xi_{t-1} : I_n = 1\}$$

with initial point  $\xi_0 = 1$ . Calculate transition probabilities (cf. [24])

$$p_{n,m} = P(\xi_{k+1} = m \mid \xi_k = n). \tag{7}$$

The first player’s reward for stopping on the  $n$ th candidate (i.e.,  $Y_n = 1$ ) is  $s_{1,n} = \frac{n}{N}$  and for continuing observations

$$c_{1,n} = \sum_{k=n+1}^N \frac{n}{k(k-1)} \frac{k}{N} = \frac{n}{N} \sum_{k=n+1}^N \frac{1}{k-1}.$$

Based on the above denotations and theory of optimal stopping, we have

$$u_{1,n} = \sup_{\tau \in \mathcal{T}_{2,n}} \mathbf{E}_n s_{1,\tau} = \max\{s_{1,n}, \mathbf{T}u_{1,n}\}. \tag{8}$$

### 2.2 Equilibrium States

Suppose that we are at some moment  $n$ , the value of the current candidate is  $x$  (seen for the Player **II**), it is relatively the best, and **both** players want to stop. If Player **II** gets the object (with probability  $1 - p$ ), his reward is  $s_{2,n}(x)$ . With probability  $p$ , Player **I** gets the object, so Player **II** must continue the observations and receive the reward  $c_{2,n}(x)$ . The situation in which in the future the opponent will find the best object is also included in the reward. A similar consideration gives us the reward for Player **I**. Let us denote the expected reward of Player **II** when he is choosing the state  $(n, x)$  which is local maximum (i.e., he effectively chooses the state  $(n, x)$ ):

$$\begin{aligned} w_{2,n}(x) &= s_{2,n}(x) - \mathbf{T}u_{2,n}(x) \quad \text{for } x \geq x_n, \text{ (cf. (6))} \\ &= s_{2,n}(x) - c_{2,n}(x) = \left( x^{N-n} - x^{N-n} \sum_{j=n+1}^N \frac{x^{j-N-1} - 1}{N - j + 1} \right) \end{aligned} \tag{9}$$

and the expected reward of Player I when he is choosing the state  $(n, x)$  which is local maximum, i.e., the relatively best at moment  $n$ :

$$\begin{aligned}
 w_{1,n} &= s_{1,n} - \mathbf{T}u_{1,n} \text{ for } n \geq n^* \text{ (cf. (8))} \\
 &= s_{1,n} - c_{1,n} = \frac{n}{N} \left( 1 - \sum_{j=n+1}^N \frac{1}{j-1} \right). \tag{10}
 \end{aligned}$$

Then the payoff matrix in the considered game is given by

$$(v_{1,n}, v_{2,n}(x)) = \begin{array}{c|cc} \mathbf{I} \backslash \mathbf{II} & S & F \\ \hline S & (2p-1)w_{1,n}; (1-2p)w_{2,n}(x) & w_{1,n}; -w_{2,n}(x) \\ \hline F & -w_{1,n}; w_{2,n}(x) & \mathbf{T}v_{1,n}, \mathbf{T}v_{2,n}(x) \end{array}, \tag{11}$$

where  $\mathbf{T}$  stands for the one-step mean operator with respect to adequate Markov process transition probability (cf. [25]).

In further analysis, comparisons of the value of payouts from the current state and the expected payouts resulting from the application of the selected strategy are performed. The technical lemma shows an example of the value of the second player’s payoff averaging operator in the event that he is interested in selecting the next potential candidate and his opponent does not interfere with it.

**Remark 1** Let us assume for that  $X_n = \max\{X_1, X_2, \dots, X_n\} = x$ . Then

$$\begin{aligned}
 \hat{w}_{2,n}(x) &= \mathbf{T}w_{2,n}(x) = \sum_{k=n+1}^N \int_x^1 w_{2,k}(y) p((n, x), (k, dy)) \\
 &= \sum_{s=1}^{N-n} s^{s-1} \int_x^1 w_{2,s+n}(y) dy = \sum_{s=1}^{N-n} s^{s-1} \tilde{w}_{2,s+n}(x), \tag{12}
 \end{aligned}$$

where

$$\begin{aligned}
 \tilde{w}_{2,j}(x) &= \frac{1-x^{N-k+1}}{N-j+1} - \sum_{r=1}^{N-j} \frac{1}{(N-j+1)(N-j-r+1)} \\
 &+ \sum_{r=1}^{N-j} \left[ \frac{x^{N-j-r+1}}{r(N-j-r+1)} - \frac{x^{N-j+1}}{r(N-j+1)} \right]. \tag{13}
 \end{aligned}$$

Since both players want to maximize their profits, we have the following conditions determining the states of observed process at which  $(S, S)$  is a Nash equilibrium:

$$\begin{cases} (2p-1)w_{1,n} \geq -w_{1,n}, \\ (1-2p)w_{2,n}(x) \geq w_{2,n}(x). \end{cases}$$



This leads to the inequalities

$$\begin{cases} \sum_{j=n+1}^N \frac{x^{j-N-1} - 1}{N - j + 1} \leq 1, \\ \sum_{j=n+1}^N \frac{1}{j - 1} \leq 1. \end{cases}$$

For player **I**, it is rational to stop when  $n \geq n^*$ , and  $n^*$  is the standard optimal threshold (cf. [4]):

$$n^* = \max \left\{ 0 \leq n \leq N : \sum_{k=n+1}^N \frac{1}{k - 1} > 1 \right\}. \tag{15}$$

Player **II**'s rational subset of states to stop are the same as in the standard one player optimal stopping problem also,

$$\mathbf{D} = \{(n, x) \in \{1, 2, \dots, N\} \times [0, 1] : X_n = x, x \geq x_n\}, \tag{16}$$

where  $x_n$  is the solution of the equation  $w_{2,n}(x) = 0$  in  $(0, 1]$  (cf. (9)),

for Player **I**,  $\tau_1 = \inf\{n > n^* : Y_n = 1\}$ ;

for Player **II**,  $\tau_2 = \inf\{n > n^* : X_n = \max\{X_1, \dots, X_n\} = x \geq x_n\}$ .

Summarizing this analysis we have

**Lemma 1** *In the game described above, the strategy  $(S, S)$  is the pure Nash equilibrium if  $X_n$  is a local maximum and  $X_n \geq x_n, n > n^*$ .*

**Remark 2** The states after which players should accept the state (i.e., keep their participation in the game) in the completed construction can be modified. We are trying to find a pair of stopping moments in the Nash equilibrium. This description is included in Corollary 1.

Suppose that  $n > n^*$  and that the value of the current observation is  $X_n = x \leq x_n$  and its relative rank is 1. If we are below this threshold, then it is more optimal for Player **II** to change his strategy on F. The best response of Player **I** to the opponent's strategy is to continue stopping if the expected future reward  $\mathbf{T}v_{1,n}$  is not greater than the actual reward  $w_{1,n}$ . The player without information knows that the opponent has more information. Since the opponent chooses strategy **F**, we know that the present value of the object is less than the threshold  $x_n$ . Suppose for a moment that we and Player **I** know this value and that it is  $x$ . If we knew this value, the future payoff would be

$$\mathbf{T}(v_{1,n} | x) = \sum_{k=n+1}^N x^{k-n-1} \left( \int_x^{x \vee x_k} \mathbb{I}_{(x, x \vee x_k]}(y) w_{1,k} dy + \int_{x \vee x_k}^1 (2p - 1) w_{1,k} dy \right),$$

where  $a \vee b = \max\{a, b\}$ ,  $\mathbb{I}_{(s,t]}(y) = 1$  when  $s < y \leq t$  and 0 otherwise, and  $w_{1,k}$  is given by (10). However, we have to average it. Knowing that the actual value is

uniformly distributed on the interval  $[0, x_n]$  (since the opponent wishes to continue the observations), we have

$$\mathbf{T}v_{1,n} = \frac{1}{x_n} \int_0^{x_n} \mathbf{T}(v_{1,n} | x) dx. \tag{17}$$

Let us consider the set  $M_1 = \{n^* < n \leq N : Tv_{1,n} \leq w_{1,n}\}$ . Note that this set is not empty. It contains the number  $\{N\}$ . Using the method of backward induction, we can find the lower bound for this set, i.e., the index  $\tilde{n} = \max\{n^* \leq n \leq N : Tv_{1,n} > w_{1,n}\}$ . The consequence of this analysis is the conclusion:

**Lemma 2** *Suppose that the current state of the process  $(n, X_n)$  is such that  $n \geq \tilde{n}$ ,  $X_n = x \leq x_n$ , and  $X_n$  is a local maximum. Then the strategy  $(S, F)$  is the pure Nash equilibrium in the game described above.*

Now suppose that  $n = \tilde{n} - 1$  and the current state of the process is  $X_{\tilde{n}-1} = x$  where  $x \leq x_n$ . Since Player I changes his strategy to  $F$ , it is necessary to check whether condition  $\mathbf{T}v_{2,n}(x) \geq w_{2,n}(x)$  is satisfied to have  $(F, F)$  the equilibrium. Indeed, it is true. Since now the reward  $w_{2,n}(x) < 0$  and the future reward is positive for  $p < 0.5$ , it is more optimal to take an action  $F$  for the player II. Now the same consideration are made for  $\tilde{n} - 2, \tilde{n} - 3, \dots$  etc. So, from this considerations, we have the conclusion.

**Lemma 3** *Suppose that the current state of the process  $(n, X_n)$  is such that  $n \leq \tilde{n}$ ,  $X_n = x \leq x_n$  and  $X_n$  is a local maximum. Then the strategy  $(F, F)$  is the pure Nash equilibrium in this state in the game described above.*

Now consider the case when  $n = n^* - 1$  and  $X_n = x > x_n$  is the local maximum. This is the opposite situation when the player II prefers to stop but the player I prefers to continue the observations without acceptance actual candidate. By analysis of the game matrix (11) at moment  $n$ , to find the subset of the state space where the strategy  $(F, S)$  is an equilibrium we compare the gain function for stopping at  $n$  and going forward (i.e., the gain function for the future). We have then

$$w_{2,n}(x) \geq \mathbf{T}v_{2,n}(x). \tag{18}$$

By (11) and the consideration of Remark 1, we have the left-hand-side of (18) as follows

$$\begin{aligned} \mathbf{T}v_{2,n}(x) = & \sum_{k=n+1}^N x^{k-n-1} \left( \int_x^{x \vee x_k} \mathbb{I}_{(x, x \vee x_k]}(y) w_{2,k}(y) dy \right. \\ & \left. + \int_{x \vee x_k}^1 (1 - 2p) w_{2,k}(y) dy \right). \end{aligned} \tag{19}$$

There is  $p$  such that the left-hand side of (18) is always bigger than the expression on the right-hand side, which is negative. Therefore, for  $n = n^* - 1$  and  $x > x_n$  it is better for player II to not change his strategy. Continuing these calculations, we get that it is also better to not change his strategy when  $n < n^*$  and  $x > x_n$ .

**Lemma 4** *Suppose that the current state of the process  $(n, X_n)$  is such that  $n < n^*$ ,  $X_n = x > x_n$  and  $X_n$  is a local maximum. Then for big enough  $p$  the strategy  $(F, S)$  is the pure Nash equilibrium in the game described above.*

**Lemma 5** *Suppose that the current state of the process  $(n, X_n)$  is such that  $n < n^*$ ,  $X_n = x \leq x_n$  and  $X_n$  is a local maximum. Then the strategy  $(F, F)$  is the pure Nash equilibrium in the game described above.*

Based on the above lemmas, we conclude by the following corollary,

**Corollary 1** *In the best choice game with the asymmetric information, there exists a Stackelberg equilibrium point, and it is given in each subgame by Lemmas 1–5.*

### 3 Numerical Example

#### 3.1 Value of the Game

The value of the game for different values of priority parameter  $p$  and  $N = 10$  is presented below.

$$\begin{aligned} (\text{val}_{1,10}, \text{val}_{2,10}) &= (0.002\ 01, 0.195\ 57), & p = 0.1, \\ (\text{val}_{1,10}, \text{val}_{2,10}) &= (0.032\ 83, 0.128\ 96), & p = 0.25, \\ (\text{val}_{1,10}, \text{val}_{2,10}) &= (0.068\ 97, 0.087\ 96), & p = e^{-1}, \\ (\text{val}_{1,10}, \text{val}_{2,10}) &= (0.136\ 62, 0.037\ 87), & p = 0.5. \end{aligned}$$

#### 3.2 Shift of the Threshold for Player I

Table 1 presents different values of the threshold  $\tilde{n}$  for different horizons and values of the priority parameter  $p$ .

**Table 1** Numbers  $\tilde{n}$

$N$	$n^*$	$\tilde{n}$					
		$p = 0.1$	$p = 0.2$	$p = 0.25$	$p = 1/3$	$p = e^{-1}$	$p = 0.5$
5	3	3	3	3	3	3	3
10	4	4	5	5	5	5	6
20	8	9	10	10	11	11	12
30	12	14	15	15	16	17	18
50	19	24	26	26	28	28	31

## 4 Conclusion

The model presented in this work was created as a fruit of reflection on real problems in the field of business and finance. In the competition between two opponents from which one of them has access to more data we have found the equilibrium states. If the priority parameter of no-information player  $p \leq 0.5$ , we have found that no-information player has to change his strategy in relation to the situation if he remained in the game alone. However, the full-information player does not intend to change his strategy. The numerical examples presented here are good presentation of the model.

It is worth adding that the importance of information in making strategic decisions when the task is dynamic and the decision maker is aware of this fact is a known research problem. In the optimal stopping problem, in the context of the role of information, it is worth mentioning the analyses in [28], or the problem of information valuation considered in the works [29, 30]. The game model considered in the work by Basu and Stettner [31] has a similar information structure. Players (agents) have partial knowledge about the state of the system, and this forces additional information filtering operations. This is additionally important because in the game under consideration, players take actions sequentially.

These examples show that further research on information modeling in multi-person decision-making processes in conjunction with modeling the psychological aspects of decision-making (v. [32]) is important both for the foundations of decision theory and for the development of probability theory methods.

**Acknowledgements** We would like to express our very great appreciation to Editors and anonymous Referees for their valuable and constructive questions, comments, and suggestions.

**Author Contributions** Marek Skarupski and Krzysztof J. Szajowski have contributed equally to this work. Both authors equally contributed to the conceptualization, methodology, formal analysis, investigation and writing—original draft preparation. M. Skarupski is responsible for the description of examples and its visualization, and K.J. Szajowski is responsible for the project conceptualization and its administration.

**Conflicts of Interest** The authors declare no conflict of interest.

**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/>.

## References

- [1] Ferguson, T.S.: Who solved the secretary problem? *Stat. Sci.* **4**(3), 282–289 (1989)
- [2] Cayley, A.: Mathematical questions with their solutions. *The Educational Times* **23**, 18–19: See the collected mathematical papers of Arthur Cayley **10**, 587–588 (1896). Cambridge Univ. Press, Cambridge (1875)

- [3] Presman, E.L., Sonin, I.M.: The best choice problem for a random number of objects. *Teor. Veroyatnost. Primenen.* **17**(4), 695–706 (1972). <https://doi.org/10.1137/1117078>
- [4] Gilbert, J.P., Mosteller, F.: Recognizing the maximum of a sequence. *J. Am. Stat. Assoc.* **61**(313), 35–73 (1966)
- [5] Dynkin, E.B.: The game variant of the optimal stopping problem. *Dokl. Akad. Nauk USSR* **185**, 241–288 (1969). Translation from *Dokl. Akad. Nauk. SSSR* **185**, 16–19 (1969)
- [6] Szajowski, K., Skarupski, M.: On multilateral incomplete information decision models. *High Freq.* **2**(3–4), 158–168 (2019)
- [7] von Stackelberg, H.F.: *Grundlagen Einer Reinen Kostentheorie*. Meilensteine Nationalökonomie, p. 131. Springer, Berlin (2009). Originally published monograph. Reprint of the 1st Edn. Wien, Verlag von Julius Springer (1932). <http://www.springerlink.com/content/978-3-540-85271-1>
- [8] Nakai, T.: Stackelberg solution for a stopping game. *J. Inf. Optim. Sci.* **18**(3), 479–491 (1997)
- [9] Ferenstein, E.Z.: Two-person non-zero-sum games with priorities. In: Ferguson, T.S., Samuels, S.M. (eds.) *Strategies for Sequential Search and Selection in Real Time*. Proceedings of the AMS-IMS-SIAM Joint Summer Research Conferences Held June 21–27, 1990. Contemporary Mathematics, vol. 125, pp. 119–133. University of Massachusetts at Amherst (1992)
- [10] Szajowski, K.: Optimal stopping of a discrete Markov process by two decision makers. *SIAM J. Control Optim.* **33**(5), 1392–1410 (1995)
- [11] Porosiński, Z., Szajowski, K.: Modified strategies in two person full-information best choice problem with imperfect observation. *Math. Jpn.* **52**(1), 103–112 (2000)
- [12] Mazalov, V.V.: *Mathematical Game Theory and Applications*, p. 414. John Wiley & Sons Ltd, Chichester (2014)
- [13] Kurano, M., Yasuda, M., Nakagami, J.: Multivariate stopping problem with a majority rule. *J. Oper. Res. Soc. Jpn.* **23**(3), 205–223 (1980)
- [14] Enns, E.G., Ferenstein, E.: The horse game. *J. Oper. Res. Soc. Jpn.* **28**(1), 51–62 (1985)
- [15] Majumdar, A.A.K.: The horse game and the OLA policy. *Indian J. Math.* **30**(3), 213–218 (1988)
- [16] Vinnichenko, S.V., Mazalov, V.V.: Games for a stopping rule of a sequence of observations of fixed length. *Kibernetika (Kiev)* **1**, 122–124 (1989)
- [17] Bruss, F.T., Drmota, M., Louchard, G.: The complete solution of the competitive rank selection problem. *Algorithmica* **22**(4), 413–447 (1998)
- [18] Szajowski, K.: Optimal stopping of a sequence of maxima over an unobservable sequence of maxima. *Zastos. Mat.* **18**(3), 359–374 (1984)
- [19] Porosiński, Z.: Optimal stopping of a random length sequence of maxima over a random barrier. *Zastos. Mat.* **20**(2), 171–184 (1990)
- [20] Neumann, P., Porosiński, Z., Szajowski, K.: On two person full-information best choice problem with imperfect observation. In: *Game Theory and Applications, II. Game Theory Applications*, pp. 47–55. Nova Science Publishers, Hauppauge, NY (1996)
- [21] Sakaguchi, M.: Optimal stopping in sampling from a bivariate distribution. *J. Oper. Res. Soc. Jpn.* **16**, 186–200 (1973)
- [22] Sakaguchi, M., Szajowski, K.: Mixed-type secretary problems on sequences of bivariate random variables. *Math. Jpn.* **51**(1), 99–111 (2000)
- [23] Chow, Y.S., Robbins, H., Siegmund, D.: *The Theory of Optimal Stopping*, p. 141. Dover Publications, Inc., New York (1991). Corrected reprint of the 1971 original
- [24] Dynkin, E.B., Yushkevich, A.A.: *Markov Processes: Theorems and Problems*, p. 237. Plenum Press, New York (1969). Translated from the Russian by James S. Wood
- [25] Shiryaev, A.N.: *Optimal Stopping Rules*. Stochastic Modelling and Applied Probability, vol. 8, p. 217. Springer, Berlin (2008). Translated from *Статистический последовательный анализ. Оптимальные правила* the 1976 Russian second edition by A. B. Aries, Reprint of the 1978 translation
- [26] Szajowski, K.: Markov stopping games with random priority. *Z. Oper. Res.* **39**(1), 69–84 (1994)
- [27] Bojdecki, T.: On optimal stopping of a sequence of independent random variables—probability maximizing approach. *Stoch. Process. Appl.* **6**(2), 153–163 (1978)
- [28] Enns, E.G.: Optimal sequential decisions with information invariance. *Sci. Math. Jpn.* **60**(3), 551–561 (2004)
- [29] Dotsenko, S.I., Marynych, A.V.: Hint, extortion, and guessing games in the best choice problem. *Cybernet. Syst. Anal.* **50**(3), 419–425 (2014)

- [30] Skarupski, M.: Full-information best choice game with hint. *Math. Methods Oper. Res.* **90**(2), 153–168 (2019)
- [31] Basu, A., Stettner, L.: Finite- and infinite-horizon Shapley games with nonsymmetric partial observation. *SIAM J. Control Optim.* **53**(6), 3584–3619 (2015)
- [32] Szajowski, K.J.: Sequential selections with minimization of failure. *J. Math. Psychol.* **111**, 102723 (2022)