EXIT GAME WITH INFORMATION EXTERNALITIES

SVETLANA BOYARCHENKO*

* Department of Economics, The University of Texas at Austin,
2225 Speedway Stop C3100, Austin, TX 78712, U.S.A.
e-mail: sboyarch@eco.utexas.edu

Preliminary and incomplete

Abstract. We analyze a two-player stopping time game with pure informational externalities. While the players are in the game, they receive deterministic revenues and incur stochastic costs. Each player incurs the cost at random times. Times of arrival of costs are modeled as Poisson processes with unknown parameters. The costs are modeled as i.i.d. random variables for each state of a continuous time Markov chain that models the evolution of the nature. The characteristics of the cost process are common to both players. As a result, players may remain active longer than a single player would do. For some observations of costs, the current state of the Markov chain cannot be uniquely inferred, therefore, there is space for a stochastic evolution of beliefs about the current state of the Markov chain and about the parameters of the Poisson process. Each player learns about the current value of the cost both when she incurs the cost, and when the other player incurs the cost. Thus, each player benefits if the other player stays in the game longer, because this increases the frequency of observations and the value of staying in the game. We demonstrate that if the players are heterogeneous, there is an equilibrium, where they exit the game sequentially, and the order of exit is determined endogenously.

The optimal stopping result is reformulated in the form of the exit region in the space of beliefs. We formulate general conditions on the evolution of beliefs near the boundary of the exit region, which ensure that the value function is smooth at a part of the boundary (absorbing part), and the condition that ensure that the value function has a kink at a part of boundary (reflecting part).

Keywords: optimal stopping, jump-diffusion process, Poisson arrival, strategic exit

JEL: C73, C61, D81

1. Introduction

1.1. **Setting and main results.** We consider a game of timing between two players who experiment with a profit generating technology, and can be either "good" or

I am thankful for comments on earlier versions of the paper to the participants of the research seminar at the University of Montreal, Canada, December 14, 2015 and International Conference on Game Theory, July 20-24, 2015, Stony Brook, NY. I am especially thankful for discussions to Sven Rady, V. Bhaskar, Caroline Thomas, Nick Klein, Dale Stahl, and Mark Feldman. The usual disclaimer applies.

"bad". If the technology is "good", it never breaks down, if the technology is "bad", costly breakdowns arrive at jump times of standard Poisson processes. The type of technology is the same for both players, so if one player observes the first breakdown, the other player knows that her technology is also "bad". The processes of breakdowns are independent but have the same characteristics so that one player can use observations of the other player to update her beliefs about the characteristics of her breakdown process. The costs of breakdowns are modeled as i.i.d. random variables for each state of a two-state continuous time Markov chain that models the evolution of the nature. The states differ by expected costs - in the "good" state, the expected cost is smaller than in the "bad" one. For some cost realizations, the current state of the nature is revealed, for others it is not. The players have common priors both about the type of technology and the state of the nature. All information, including the players' payoffs is public, therefore the players have the same beliefs as time evolves. At each point in time, each player can make an irreversible decision to stop experimentation and exit the game. The value of an outside option is zero.

An interesting feature of our model that makes it different from a typical setup in the part of literature dealing with strategic learning and experimentation known as exponential bandits literature, is that the beliefs are two-dimensional before the first breakdown happens. The longer no breakdown is observed, the higher is the belief that the technology is "good" no matter what the initial prior was. The beliefs about the state of the nature may move in the same direction in the sense that the beliefs that the state of the nature is "good" increase, but they can also move in the opposite direction. The dynamics of these beliefs depends on the initial prior assigned to the probability of the "good" state. Namely, if the prior is below the long run probability of the "good" state, the posterior beliefs are increasing in time (prior to any breakdown). If this is the case, none of the players will stop before the first breakdown happens. However, if the initial prior is above the long run probability of the "good" state, the posterior beliefs are decreasing in time. When beliefs about the state of the nature and those about the quality of the technology evolve in the opposite directions, it may become optimal for one or both players to exit the game before any breakdown happens, as our model demonstrates. Such outcome is impossible in a model with conclusive breakdowns, e.g., in Keller and Rady [27]. More generally, the optimally stopping strategy in exponential bandits literature is characterized in terms of a cutoff belief - as soon as the cutoff level is reached, the experimentation stops. With two-dimensional beliefs, it is impossible to talk about a cutoff belief, and the stopping region is not an interval but an area in a unit square. After the first breakdown, the rate of arrival of the costs becomes known, but the players are still facing uncertainty of the cost process, therefore they do not necessarily exit immediately after the first breakdown. Instead, they recognize the option value of waiting and stop experimenting the first time their beliefs about the probability of the "good" state enter the stopping region. We demonstrate that the stopping region if of the form $[0, p^*] \in [0, 1]$, where p^* is the cutoff belief. It may be the case that $p^* = 0$, which means that a player never exits before the next

breakdown happens. Depending on the parameters of the model, the cutoff belief may be higher than the long run probability of the "good" state - in this case, the player may exit between two consecutive ; or the cutoff belief may be higher than the long run probability of the "good" state - in this case, the player may exit only at the time of a breakdown. In the richer two-dimensional beliefs setting, exit prior to the first break down or exit at the time of the first breakdown (or later) may happen for the same set of parameters of the model.

The biggest challenge of the model is the fact that the players' strategies are stopping times, and the filtration depends on the number of active players. So, while both players are active, the filtration is relatively fine, and the strategies are stopping times w.r.t. this fine filtration. As soon as one of the players quits, the remaining player faces a relatively coarse filtration, and her strategy is the stopping time with respect to this coarse filtration. To deal with these difficulties, it becomes necessary to impose additional restrictions on admissible strategies.

As the first step, we consider the case when one player gets lower revenue flow than the other one and show that the latter player will never exit earlier than the former one. Following the standard terminology, we call the leader the player who quits first, and the follower is the other player. We prove that the leader's and the follower's problems have solutions and characterize these solutions in terms of robust recursive procedures. We also provide sufficient conditions on parameters of the model that guarantee exit in finite time.

In models with strategic learning, players impose information externalities on their counterparts as long as they keep experimenting. Even when all the players stop at an individually optimal cutoff belief, i.e., equilibria exhibit no encouragement effect, the time it takes for the player's beliefs to reach the cutoff level depends on the presence of the other player(s). We demonstrate that the cutoff belief of a single player is higher than the cutoff belief of the same player, when the other player is still active, provided the information externality is positive. This differs from a result in Heidhues et al. [21], where in the case of public payoffs all the players stop experimenting at the individually optimal level of beliefs. In the op.cit., the result is driven by free riding incentives. In our model, there are no incentives to free ride. We leave for the future work the case when exit is not irreversible, so a player may have an incentive to quit for some time and re-enter when the beliefs about the "good" state become high enough. In this scenario, there may be a free riding problem.

If the information externality is always positive (i.e., the value of a single player is smaller than the value of the same player when (s)he observes breakdowns more frequently), then it may be possible to argue the existence of a symmetric equilibrium, when the players are symmetric. Unfortunately, we were not able to prove that the information externality in our model is always positive, though we observed this effect in all numerical experiments. If this is indeed the case, there is a role for a policy maker to provide incentives for a player who wants to exit earlier than the other one(s) to stay longer.

In the optimal stopping literature, and, recently, in the literature on learning and strategic experimentation, popular topics are the study of the regularity of the value function and early exercise boundary if the latter is time dependent. In diffusion models in continuous time, the value functions of perpetual American options with standard payoffs and real options are smooth (one says that *the smooth pasting principle* holds at the boundary of the inaction region). The same holds for American options with finite time horizon, at each time moment before expiry. See [38] for details. For perpetual American options and real options in discrete time, the smooth pasting principle fails [10, 13, 14]. The smooth pasting principle fails in Lévy models of finite variation with drift pointing from the boundary [12, 11, 14]. In cases when the smooth pasting principle fails, it was proved that the *continuous pasting principle* was sufficient to find the exercise threshold. In the aforementioned papers, the failure of the smooth pasting principle was explained in terms of the characteristic function of the infimum process evaluated at exponentially distributed random time independent of the process. In [1], an equivalent explanation in terms of the atom of the same distribution was given. In sequential testing problems for Poisson processes, conditions for the failure of the smooth pasting principle and validity of the continuous pasting principle were established in [38], and, in problems of strategic exploration and experimentation, in [28] and [35]. Keller and Rady [28] also notice an interesting difference in the strategic experimentation in the models with breakthroughs ([27]) and with breakdowns ([28]). In the former model, the smooth pasting principle holds, and, in the latter model, the smooth pasting principle does not hold. In a model with costly breakdowns incurred at jump times of a Poisson process, where the costs follow a Lévy process, Boyarchenko and Levendorskiǐ [15] formulate the following dichotomy: either the optimal exercise policy formulated in terms of the exit time is regular at the boundary of the inaction region at the moment of the last observation and the smooth pasting principle holds, or the optimal exercise policy has a jump, and the value function has a kink at the boundary of the inaction region.

In the current paper, we also study the regularity of the value function and the optimal exercise strategy. After the first breakdown happened, and the stopping region is characterized in terms of a single cutoff belief, the value function is smooth at the exercise boundary when the cutoff belief is higher than the long run probability of the "good" state, and the optimal exercise policy is regular at the boundary. If the cutoff belief is lower than the long run probability of the "good" state, the value function has a kink at the boundary, and the optimal exercise policy has a jump. This differs from the results in Keller and Rady [27, 28].

The situation becomes more interesting before the first breakdown happens. We show that the value function is smooth (and the exercise policy is regular) if a trajectory of the corresponding beliefs about the quality of the technology and state of the nature tends towards the boundary of the stopping region. If a trajectory of the corresponding beliefs about the quality of the technology and state of the nature moves away from the boundary of the stopping region, the value function has a kink at the boundary (and the optimal exercise policy has a jump).

1.2. **Literature review.** There is an excellent and comprehensive review of the state of the art in learning, experimentation and information design provided by Hörner and Skrzypacz [22].

A typical setting in exponential bandits models involves several players facing identical copies of the same slot machine. Bandit problems are used to study the trade-off between exploration and exploitation. Bandit models were successfully used in various settings in economics, for example, learning and matching in labor markets, monopolist pricing with unknown demand, choice between R&D projects, or financing of innovations (see, e.g., [3, 4, 5, 6, 7, 8, 26, 39, 41, 42] and references therein). The situation becomes even more interesting if several DM's participate in experimentation because in this case information externalities are present, and, potentially, free riding problems may arise. See, for example, [9, 18, 24, 25, 27, 28].

In models based on two-armed bandits, the decision maker (DM) has to decide on optimal allocation of her time between the safe action (arm of the bandit) and the risky action (arm). The safe action generates a payoff given by a known distribution (in many instances, this payoff is deterministic), and the risky action generates an unknown payoff. In continuous time models, the latter payoff follows a certain continuous time stochastic process whose parameters are not known. For example, Bolton and Harris [9] model the unknown payoff as a Brownian motion with unknown drift and known variance in a model of strategic experimentation. Decamps et al. [19, 20] study timing a fixed size investment into a risky project with the payoff generated by a Brownian motion with unknown drift and known variance. Keller et al. [25], Keller and Rady [27, 28] use a Poisson process with unknown rate of arrival to model the risky arm. Decamps and Mariotti [18] study a duopoly model of investment where a signal about the quality of the project is modeled as a Poisson process. Cohen and Solan [16] bridge the gap between the Brownian motion and Poisson bandits and consider two-armed bandits, where the risky arm yields stochastic payoffs generated by a Lévy process. In two-armed bandit problems, it is typically assumed that the expected payoff generated by the risky arm is higher (respectively, lower) than the expected payoff generated by the safe arm, if the risky arm is, in some sense, "good" (respectively, "bad"). The optimal stopping rule in such problems is of a cut-off type: as long as the posterior belief that the payoff distribution of the risky arm is "good" has not reached the cut-off, the DM continues experimentation with the risky arm. When the posterior belief reaches the cut-off, the DM switches to the safe arm. In particular, in learning models with conclusive breakthroughs or breakdowns, this means that the first observation of a success or a failure implies immediate action (for example, adoption or discarding of a risky technology). Recent developments include (but are not limited to) correlated risky arms as in Klein and Rady [29] and Rosenberg et al. [40], or private payoffs as in Heidhues et al. [21] and Rosenberg et al. [40], or departures from Markovian strategies as in H¨orner et al. [24]. For other developments see Hörner and Skrzypacz [22] and references there in.

The papers which are mostly close to our paper are Keller and Rady [27, 28] and Rosenberg et al. [40]. Keller and Rady [28] study the case of costly breakdowns that

arrive at the jump times of Poisson processes which are independent. Rosenberg et al. [40] consider an irreversible exit problem in a model with breakthroughs with correlated risky arms both in the case when payoffs are public and private.

The settings of the above models ignore such features as dynamics of costs of breakdowns or profits of breakthroughs. For example, nuclear stations may have easily repairable breakdowns, or breakdowns on the scale of Chernobyl. When making decisions about optimal abandonment of a project with costly breakdowns, one takes into account not only the frequency of failures, but also the cost incurred after each breakdown, and whether the manufacturer successfully (say, a number of cars recalled decreases) or unsuccessfully (say, a new version of software fails no matter what) improves the equipment. Similarly, when a breakthrough happens, the profitability of the project may depend on specific market conditions (e.g., new shale oil extraction technology may be more or less profitable depending on oil prices).

1.3. **Empirical facts.** Here are some examples that demonstrate that not only the occurrence of a costly event, but also the size of the loss matters, and actions not necessarily take place right after an observation of a random cost.

1. Before the 9/11 terrorist attacks, the terrorism risk was included as an unnamed peril into commercial insurance contracts in the U.S. Such events as the first bombing of the World Trade Center in New York in 1993, or the 1998 bomb attacks on the U.S. embassy in Nairobi, Kenya, which caused significant insurance losses (see [30] for details) did not change the attitude of the U.S. insurers or international reinsurers to terrorism coverage. The situation changed dramatically after 9/11 attacks, which resulted in unprecedented losses. Private reinsurers, who covered the majority of these losses exited the market, and a few months after the attacks, the insurers excluded terrorism from their policies in most states (see [30] for details).

2. Other examples of how insurers who suffered large losses from a disaster are reluctant to continue offering coverage against this risk are hurricane insurance market in Florida and earthquake insurance market in California. Even though hurricanes in Florida are not rare events, the especially large losses during 2004 and 2005 hurricane seasons caused a failure of private insurance market. In California, private insurance companies decided to stop covering the residential property after the Northridge earthquake of 1994. Moreover, the earthquake happened in January 1994, but private insurance companies decided to quit the residential property market only in 1995. Hence, a one year delay was observed.

3. Following the shortage of reinsurance after such catastrophic events as Hurricane Andrew in 1992 and Northridge earthquake in 1994, there emerged a special financial contracts in order to complement reinsurance in covering large losses. One of such contracts are catastrophe (CAT) bonds. All CATs are structured to pay on triggers; in particular, there are CAT bonds that pay on insurer-specific catastrophe losses, or insurance-industry catastrophe loss indices (see, e.g., [17] for details).

4. Before the Fukushima 2011 disaster, 442 nuclear power reactors in 30 countries produces 14 per cent of all world's electricity. This number dropped to 11 per cent in 2012 as 15 reactors, mainly in Germany and Japan, exited service.

1.4. **The structure of the paper.** The rest of the paper is organized as follows. In Section 2, we set up the model, describe the evolution of beliefs, give definitions of strategies of the players and define a Nash equilibrium of the game. In Section 3, the main steps of the solution for the case of asymmetric players are presented. For the case when at least one breakdown had happened, we present a convenient recursive procedure which calculates simultaneously the optimal stopping strategy and value function. If no breakdown was observed, integral representations of the value functions are no longer recursive, but they involve value functions which were calculated for the case when at least one breakdown had happened. Section 4 studies the regularity of the value functions. Section 5 concludes.

2. The model

2.1. **The setup.** We consider the game of timing, characterized by the following structure. Time $t \in \mathbb{R}_+$ is continuous, and the discount rate is $r > 0$. Two players experiment with technologies that generate a deterministic payoff stream. An active player j gets a constant revenue stream $R_j > 0$. Assume that $R_2 \geq R_1$. The technologies may experience costly breakdowns. The breakdowns arrive at the jump times of standard Poisson processes N^1 and N^2, whose intensity depends on an unknown type of the technology $\theta \in \{0, 1\}$. If $\theta = 1$ ("bad" technology) the intensity is $\lambda' > 0$ for both players (this assumption can be relaxed, and the players may have different intensities); if $\theta = 0$ ("good" technology), the intensity is 0. Thus, in this paper, we focus on so called conclusive breakdowns, which means that the "good" technology never breaks down. If the technology is "bad" player j incurs a costly breakdown at random time τ_j . Assume that the processes of breakdowns are independent but have the same characteristics so that one player can use observations of the other player to update her beliefs about the characteristics of her breakdown process. We leave for the future study the case when the players can have technologies of different types, and the types may be positively of negatively correlated as in Rosenberg et al. [40].

The initial common prior π_0 assigns the probability $\pi_0 \in [0, 1]$ to $\lambda = 0$.

If the technology is "bad", the cost of breakdown follows a stochastic process $C = \{c_t\}_{t \geq 0}$ independent of the Poisson processes N^1 and N^2. At each t, a random variable c_t can take one of three possible values: $c_t \in \{c_l, c_m, c_h\}$, where numbers $c_h > c_m > c_l > 0$ are publicly known. The distribution of c_t depends on the state of nature at time t. The state at each time t may be "good" (state 1) or "bad" (state 2). The distribution of c_t conditioned on the state realization at time t is specified as follows:

$$\mathrm{Prob}(c_t = c_h |\, g) = 0,\ \mathrm{Prob}(c_t = c_m |\, g) = \alpha_1,\ \mathrm{Prob}(c_t = c_l |\, g) = 1 - \alpha_1,$$
$$\mathrm{Prob}(c_t = c_h |\, b) = 1 - \alpha_2,\ \mathrm{Prob}(c_t = c_m |\, b) = \alpha_2,\ \mathrm{Prob}(c_t = c_l |\, b) = 0.$$

Thus, if c_h or c_l are observed, the state of the nature is known for sure. If c_m is observed, the state of the nature is not revealed. The states of the nature follow a continuous time Markov chain (CTMC) with the transition rate matrix

$$L = \begin{pmatrix} -\mu_1 & \mu_1 \\ \mu_2 & -\mu_2 \end{pmatrix}, \quad \mu_1 > 0,\, \mu_2 > 0.$$

The probabilities α_1 and α_2 and transition rates μ_1 and μ_2 are known publicly and do not change over time. The initial common prior assigns probability $\bar{p}_0 \in [0,1]$ to state 1.

2.2. **Filtered measurable spaces.** The evolution of beliefs (see explicit formulas in Section 2.3 below) define the probability measures on the following measurable spaces that naturally arise in the model.

I. $(\Omega^i, \mathcal{F}^i, \{\mathcal{F}^i_t\}_{t=0}^\infty)$, generated by the cost process of player i, $i = 1,2$, under the assumption that this player keeps experimenting forever.

II. $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^\infty)$, generated by the cost processes of both players, under the assumption that they keep experimenting forever.

To be more specific, the generic element of Ω^i can be identified with a sequence $\{\tau_i^k, c_{\tau_i^k}\}_{k=1}^\infty$, where $0 < \tau_i^1 < \tau_i^2 < \cdots$ are the jump times of N^i, and $c_{\tau_i^k}$ are costs incurred by player i at these times. The generic element of Ω can be identified with a sequence of times and costs incurred by each of the players. Denote by ω^i the generic element of Ω^i and by ω the generic element of Ω, and by $\mathcal{M}^i$ and $\mathcal{M}$ the set of stopping times w.r.t. to $\mathcal{F}^i$ and $\mathcal{F}$, respectively.

The initial beliefs $(\pi_0, \bar{p}_0)$ give rise to the probability measures $\mathbb{P}^i_{\pi_0,\bar{p}_0}$ and $\mathbb{P}_{\pi_0,\bar{p}_0}$ on $(\Omega^i, \mathcal{F}^i, \{\mathcal{F}^i_t\}_{t=0}^\infty)$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^\infty)$, respectively. Let $\mathbb{E}^i_{\pi_0,\bar{p}_0}$ and $\mathbb{E}_{\pi_0,\bar{p}_0}$ be the corresponding expectation operators.

2.3. **Evolution of beliefs.** Consider the evolution of beliefs until the first cost is incurred at random time τ. Let $\pi(t)$ be a posterior belief about the event $\lambda = 0$. Assuming $\pi(t)$ is known, and the disaster does not happen during the next infinitesimally small time interval Δt, the updated belief is

$$\pi(t + \Delta t) \quad = \quad \mathrm{Prob}\,(\lambda = 0 \mid \tau > t + \Delta t) = \frac{\pi(t)}{\pi(t) + (1 - \pi(t))e^{-\lambda' \Delta t}} + o(\Delta t).$$

Calculating the derivative of the RHS w.r.t. Δt at $\Delta t = 0$ we derive the ordinary differential equation for the evolution of beliefs: $(\pi)'(t) = \lambda' \pi(t)(1 - \pi(t))$. The solution is standard and well-known: if $\pi(0) = \pi_0$, then, until the moment of the first observation, π and λ evolve according to

$$\pi(\pi_0; t) \quad = \quad \frac{\pi_0}{(1 - \pi_0) \cdot e^{-\lambda' t} + \bar{\pi}_0}, \tag{2.1}$$

$$\lambda(\pi_0; t) \quad = \quad \lambda' - \frac{\pi_0 \lambda'}{(1 - \pi_0) \cdot e^{-\lambda' t} + \pi_0}. \tag{2.2}$$

After the first observation, $\pi(\pi_0, t) = 0$, and $\lambda(\pi_0, t) = \lambda'$. We introduce the point processes $\tilde{N}^i_{\pi_0}(t)$ with the stochastic intensity $\lambda(\pi_0, t)$.

Let $p_{ij}(t)$ $(i, j = 1, 2)$ denote the probability of being in state j at time t conditioned on being in state i at time 0. Standard calculations show that

$$p_{11} = \frac{\mu_2}{\mu_1+\mu_2} + \frac{\mu_1}{\mu_1+\mu_2} e^{-(\mu_1+\mu_2)t}, \tag{2.3}$$

$$p_{12} = \frac{\mu_1}{\mu_1+\mu_2} - \frac{\mu_1}{\mu_1+\mu_2} e^{-(\mu_1+\mu_2)t}, \tag{2.4}$$

$$p_{21} = \frac{\mu_2}{\mu_1+\mu_2} - \frac{\mu_2}{\mu_1+\mu_2} e^{-(\mu_1+\mu_2)t}, \tag{2.5}$$

$$p_{22} = \frac{\mu_1}{\mu_1+\mu_2} + \frac{\mu_2}{\mu_1+\mu_2} e^{-(\mu_1+\mu_2)t}. \tag{2.6}$$

Prior to any observation, the beliefs about the state of nature evolve as follows:

$$\begin{aligned} p_j(\bar{p}_0, t) &= \bar{p}_0 p_{1j}(t) + (1-\bar{p}_0) p_{2j}(t) \\ &= p_{2j}(t) + \bar{p}_0 (p_{1j}(t) - p_{2j}(t)), \end{aligned} \tag{2.7}$$

where $p_{ij}(t)$ are given by (2.3) - (2.6). Using (2.3) - (2.6), we derive

$$\begin{aligned} p_1(\bar{p}_0, t) &= p_{21}(t) + \bar{p}_0 e^{-(\mu_1+\mu_2)t} \\ &= \frac{\mu_2}{\mu_1+\mu_2} + \left(\bar{p}_0 - \frac{\mu_2}{\mu_1+\mu_2}\right) e^{-(\mu_1+\mu_2)t}. \end{aligned} \tag{2.8}$$

If a cost c_τ is incurred at time τ, beliefs about the "good" state are updated according to the Bayes rule. Namely,

$$p_{1u}(\bar{p}_0, \tau |\, c_\tau) = \begin{cases} 0, & \text{if } c_\tau = c_h, \\ \frac{\alpha_1 p_1(\bar{p}_0, \tau-)}{\alpha_2 + (\alpha_1-\alpha_2) p_1(\bar{p}_0, \tau-)}, & \text{if } c_\tau = c_m, \\ 1, & \text{if } c_\tau = c_l. \end{cases} \tag{2.9}$$

Notice that if c_h is observed at time τ, the beliefs about the probability of "good" state at time τ jump down to zero. If c_l is observed at time τ, the beliefs about the probability of "good" state at time τ jump up to one. If c_m is observed at time τ, the beliefs about the probability of "good" state at time τ jump up iff $\alpha_1 > \alpha_2$, i.e., the probability of incurring c_m in the "good" state is higher than in "bad" state as the result below shows.

Claim 2.1. *(i) The updated belief about the probability of "good" state at time τ is non-decreasing in $\bar{p}_1^0$ (strictly increasing if $c_\tau = c_m$).*
(ii) The updated belief about the probability of "good" state at time τ conditioned on $c_\tau = c_m$ jumps up (respectively, down) if $\alpha_1 > \alpha_2$ (respectively, $\alpha_1 < \alpha_2$).

Proof. (i) Evidently, $p_{1u}(\bar{p}_0, t|\, c_\tau)$ is independent of $\bar{p}_0$ if $c_\tau = c_h$ or $c_\tau = c_l$. To show that $p_{1u}(\bar{p}_0, t|\, c_\tau = c_m)$ is strictly increasing in $\bar{p}_0$, calculate the derivative

$$\begin{aligned}
&\frac{\partial p_{1u}(\bar{p}_0, \tau|\, c_\tau = c_m)}{\partial \bar{p}_0} \\
&= \frac{\alpha_1((\alpha_1-\alpha_2)p_1(\bar{p}_0,\tau)+\alpha_2) - \alpha_1(\alpha_1-\alpha_2)p_1(\bar{p}_0,\tau)}{((\alpha_1-\alpha_2)p_1(\bar{p}_0,\tau)+\alpha_2)^2} \cdot \frac{\partial p_1(\bar{p}_0,\tau)}{\partial \bar{p}_0} \\
&= \frac{\alpha_1\alpha_2 e^{-(\mu_1+\mu_2)\tau}}{((\alpha_1-\alpha_2)p_1(\bar{p}_0,\tau)+\alpha_2)^2} > 0.
\end{aligned}$$

(ii)

$$\begin{aligned}
p_{1u}(\bar{p}_0, t|\, c_\tau = c_m) - p_1(\bar{p}_0,\tau) &= \frac{(\alpha_1-\alpha_2)p_1(\bar{p}_0,\tau)(1-p_1(\bar{p}_0,\tau))}{\lambda_2 + (\alpha_1-\alpha_2)p_1((\bar{p}_0,\tau)}; \\
p_{1u}((\bar{p}_0, t|\, c_\tau = c_m) - p_1((\bar{p}_0,\tau) &> 0 \;\Leftrightarrow\; \alpha_1 > \alpha_2.
\end{aligned}$$

□

For each moment of observation τ^k of a costly breakdown $k = 1, 2, \ldots$, define the posterior belief $\bar{p}_k$ as

$$\bar{p}_k(\bar{p}_{k-1}, \tau^k|\, c_{\tau^k}) = \begin{cases} 0, & \text{if } c_{\tau^k} = c_h, \\ \frac{\alpha_1 p_1(\bar{p}_{k-1},\tau^k-)}{\alpha_2+(\alpha_1-\alpha_2)p_1(\bar{p}_{k-1},\tau^k-)}, & \text{if } c_{\tau^k} = c_m, \\ 1, & \text{if } c_{\tau^k} = c_l, \end{cases} \tag{2.10}$$

where for each $t \in (\tau^k, \tau^{k+1}]$,

$$p_1(\bar{p}_k, t) = \frac{\mu_2}{\mu_1+\mu_2} + \left(\bar{p}_k - \frac{\mu_2}{\mu_1+\mu_2}\right) e^{-(\mu_1+\mu_2)(t-\tau^k)}. \tag{2.11}$$

The evolution of the random variable c_t can be described as follows: for $s \in \{l, m, h\}$, let $p_s(t)$ be the probability of cost c_s at time $t \in [0, \tau^1]$ before the first breakdown, or $t \in (\tau^k, \tau^{k+1}]$

$$p_l(\bar{p}_k, t) = (1-\alpha_1)p_1(\bar{p}_k, t), \tag{2.12}$$

$$p_m(\bar{p}_k, t) = \lambda_2 + (\alpha_1-\alpha_2)p_1(\bar{p}_k, t), \tag{2.13}$$

$$p_l(\bar{p}_k, t) = (1-\alpha_2)(1-p_1(\bar{p}_k, t)). \tag{2.14}$$

2.4. **Histories and actions.** At each point in time player i can make an irreversible stopping decision, conditional of the history of the game. The value of an outside option in case of ext is zero. In the current setting, we consider the case when all payoffs and the players' actions are public information. At any $t \geq 0$, the history of the game includes observations of all costly breakdowns (including the empty set if not costs were incurred by either player up to time t) and the actions of the players. As far as the actions are concerned, only two sorts of histories matter in the stopping game: (i) both players are still in the game; (ii) at least one player exited the game.

Let $t_i \in \mathbb{R}_+$ denote the actual exit time of player i. Define the function

$$\tilde{t}_i(t) = \begin{cases} t_i, & \text{if } t_i \leq t, \\ \infty, & \text{otherwise.} \end{cases}$$

Let τ_i^s denote a random time, when a cost was incurred by player i for the s^{th} time, when player i was active. The history of observations at any $t \geq 0$ is

$$O_t = \{c(\tau_1^{s'})\}_{s' \leq t \wedge t_1} \cup \{c(\tau_2^{s''})\}_{s'' \leq t \wedge t_2} \cup \{\emptyset\}.$$

Thus, a typical history at time t is $h_t(O_t, \tilde{t}_1(t), \tilde{t}_2(t))$.

In the future, we will consider the case when player i can only observe that player j had a breakdown and whether player j exited or not after the breakdown. The cost incurred will be private information. This setting will make informational interactions more complicated.

2.5. **Strategy space and equilibrium concepts.** Since all the information is public, players have identical beliefs. Notice that before the first breakdown, the evolution of posterior beliefs $(\pi(t), p_1(\bar{p}_0, t))$ is deterministic. The longer nothing happens, the higher is the belief assigned to $\lambda = 0$. The beliefs about the state of the nature can move in the same or in the opposite direction compared to beliefs about the quality of technology. Indeed, if $\bar{p}_0 \leq \mu_2/(\mu_1+\mu_2)$, which is the long-run probability of state 1, then the posterior belief about the probability of the "good" state is increasing in time or remains constant, therefore, none of the players will exit until the first breakdown happens. On the contrary, if $\bar{p}_0 < \mu_2/(\mu_1 + \mu_2)$, the posterior belief $p_1(\bar{p}_0$ is decreasing in time. Hence, the longer nothing happens, the higher is the belief that the quality of technology is "good", and the lower is the belief that the state of nature is "good." Therefore, is some cases and agent may exit even before any breakdown happens. This effect is not observed in the standard exponential bandits models with conclusive breakdowns.

Definition 2.2. A strategy for player $i \in 1, 2$ in the game starting at $(\pi_0, \bar{p}_0) \in [0, 1] \times [0, 1]$ is a real valued function

$$\tilde{T}_i^{\pi_0,\bar{p}_0} : \Omega \ni \omega \mapsto \tilde{T}_i^{\pi_0,\bar{p}_0}(\omega) \in [0, \infty]. \tag{2.15}$$

which is left continuous with right limits a.s.

Consider a sample path $\omega \in \Omega$. If $\tilde{T}_i^{\pi_0,\bar{p}_0}(\omega) = 0$, player i exits immediately. If $0 < \tilde{T}_i^{\pi_0,\bar{p}_0}(\omega) < \infty$, player i exits at $\tilde{T}_i^{\pi_0,\bar{p}_0}(\omega)$. If $\tilde{T}_i^{\pi_0,\bar{p}_0}(\omega) = \infty$, player i never exits. If

$$\tilde{T}_i^{\pi_0,\bar{p}_0}(\omega) \leq \tilde{T}_j^{\pi_0,\bar{p}_0}(\omega),$$

player i is the leader, and player j is the follower.

If player i is the only one remaining in the game, $\tilde{T}_i^{\pi_0,\bar{p}_0} \in \mathcal{M}^i$, where $\mathcal{M}^i$ is the class of stopping times with respect to the filtration $\mathcal{F}^i$. If player $j \neq i$ pre-commits not to exit earlier that player i, $\tilde{T}_i^{\pi_0,\bar{p}_0} \in \mathcal{M}$, where $\mathcal{M}$ is the class of stopping times with respect to the filtration $\mathcal{F}$.

Consider a history h_t s.t.

$$O_t = \{\emptyset\}, \tilde{t}_j(t) = t_j, \tilde{t}_i(t) = \infty. \tag{2.16}$$

Let $\tilde{S}_i(\pi_0, \bar{p}_0)$ denote the value of player i after the history (2.16). Then

$$\tilde{S}_i(\pi_0, \bar{p}_0) = \sup_{\tau \in \mathcal{M}^i} \mathbb{E}^i_{\pi_0, \bar{p}_0} \left[\int_0^\tau e^{-rt} \left(R_i \, dt - c_t \, d\tilde{N}^i_{\pi_0}(t) \right) \right] \tag{2.17}$$

Consider a history h_t s.t.

$$O_t = \{\emptyset\}, \tilde{t}_i(t) = \tilde{t}_j(t) = \infty. \tag{2.18}$$

Let $\tilde{L}_i(\pi_0, \bar{p}_0)$ denote the value of player i after the history (2.18) in the case when player i is the leader, and j pre-commits to be the follower and exit no earlier than i. Then

$$\tilde{L}_i(\pi_0, \bar{p}_0) = \sup_{\tau \in \mathcal{M}} \mathbb{E}_{\pi_0, \bar{p}_0} \left[\int_0^\tau e^{-rt} \left(R_i \, dt - c_t \, d\tilde{N}^i_{\pi_0}(t) \right) \right] \tag{2.19}$$

Note that although the integrands in (2.17) and (2.19) are the same, the filtered probability spaces and expectation operators are different.

Let $\tilde{F}_i(\pi_0, \bar{p}_0, T_j)$ denote the value of player i after the history (2.18) in the case when j is the leader, who exits at T_j, and i pre-commits to be follower. Then

$$\begin{aligned} \tilde{F}_i(\pi_0, \bar{p}_0, T_j) \quad = \quad & \sup_{\tau \in \mathcal{M}^i, \tau \geq T_j} \left\{ \mathbb{E}_{\pi_0, \bar{p}_0} \left[\int_0^{T_j} e^{-rt} \left(R_i \, dt - c_t \, d\tilde{N}^i_{\pi_0}(t) \right) \right] \right. \\ & \left. + \mathbb{E}^i_{\pi_0, \bar{p}_0} \left[\int_{T_j}^\tau e^{-rt} \left(R_i \, dt - c_t \, d\tilde{N}^i_{\pi_0}(t) \right) \right] \right\} \end{aligned} \tag{2.20}$$

If there is no pre-commitment, the definition of classes of admissible stopping times, which can lead to Nash equilibrium, is more involved. We give the following definitions.

Definition 2.3. We say that the exit time $\tilde{T}_i^{\pi_0, \bar{p}_0}$ of player i is admissible if $\tilde{T}_i^{\pi_0, \bar{p}_0}$ is stopping time w.r.t. the filtration $\mathcal{F}^i$, and there exist a stopping time $\tilde{T}_j^{\pi_0, \bar{p}_0}$ w.r.t. the filtration $\mathcal{F}^j$ such that $\tilde{T}_i^{\pi_0, \bar{p}_0} \wedge \tilde{T}_j^{\pi_0, \bar{p}_0}$ is a stopping time w.r.t. the filtration $\mathcal{F}$.

If these conditions are satisfied, we write $\tilde{T}_i^{\pi_0, \bar{p}_0} \in \mathcal{M}^i_0$.

Definition 2.4. We say that the exit time $\tilde{T}_i^{\pi_0, \bar{p}_0}$ of player i is admissible for a given $\tilde{T}_j^{\pi_0, \bar{p}_0} \in \mathcal{M}^j_0$, if $\tilde{T}_i^{\pi_0, \bar{p}_0}$ is a stopping time w.r.t. the filtration $\mathcal{F}^i$, and $\tilde{T}_i^{\pi_0, \bar{p}_0} \wedge \tilde{T}_j^{\pi_0, \bar{p}_0}$ is a stopping time w.r.t. the filtration $\mathcal{F}$.

If these conditions are satisfied, we write $\tilde{T}_i^{\pi_0, \bar{p}_0} \in \mathcal{M}^i(\tilde{T}_j^{\pi_0, \bar{p}_0})$.

Definition 2.5. Given $\tilde{T}_j^{\pi_0,\bar{p}_0} \in \mathcal{M}_0^j$, and initial beliefs $(\pi_0, \bar{p}_0)$, player $i \neq j$ solves the optimization problem

$$\begin{aligned} V_i(\pi_0, \bar{p}_0, \tilde{T}_j^{\pi_0,\bar{p}_0}) \quad &= \sup_{\tau \in \mathcal{M}^i(\tilde{T}_j^{\pi_0,\bar{p}_0})} \left\{ \mathbb{E}_{\pi_0,\bar{p}_0} \left[\int_0^{\tau \wedge \tilde{T}_j^{\pi_0,\bar{p}_0}} e^{-rt} \left(R_i\, dt - c_t\, d\tilde{N}^i_{\pi_0}(t) \right) \right] \right. \\ &\left. + \mathbb{E}^i_{\pi_0,\bar{p}_0} \left[\int_{\tau \wedge \tilde{T}_j^{\pi_0,\bar{p}_0}}^{\tau} e^{-rt} \left(R_i\, dt - c_t\, d\tilde{N}^i_{\pi_0}(t) \right) \right] \right\}. \end{aligned} \tag{2.21}$$

Denote by $\mathcal{T}_i(\pi_0, \bar{p}_0, \tilde{T}_j^{\pi_0,\bar{p}_0})$ a maximizer.

Definition 2.6. A pair of strategies $(\tilde{T}_1^{\pi_0,\bar{p}_0}, \tilde{T}_2^{\pi_0,\bar{p}_0})$ is a Nash equilibrium for the game, starting at the beliefs level $(\pi_0, \bar{p}_0)$ if, for each pair $(i,j) \in \{(1,2),(2,1)\}$, $\tilde{T}_i^{\pi_0,\bar{p}_0} \in \mathcal{M}_0^i$, $\tilde{T}_j^{\pi_0,\bar{p}_0} \in \mathcal{M}^j(\tilde{T}_i^{\pi_0,\bar{p}_0})$, and $\mathcal{T}_i(\pi_0, \bar{p}_0, \tilde{T}_j^{\pi_0,\bar{p}_0}) = \tilde{T}_i^{\pi_0,\bar{p}_0}$.

For any time $t > 0$, define a proper subgame as a timing game that starts at the decision node $(0, \bar{p}_k)$ $(k = 1, 2, \ldots)$.

Definition 2.7. A pair of strategies is a subgame perfect equilibrium, if for any $k = 1, 2, \ldots$, $(\tilde{T}_1^{0,\bar{p}_k}, \tilde{T}_2^{0,\bar{p}_k})$ is a Nash equilibrium.

For $k \geq 1$, denote $T_i^{\bar{p}_k} := \tilde{T}_i^{0,\bar{p}_k}$.

Consider a history h_t s.t.

$$\#O_t = k, \tilde{t}_j(t) = t_j, \tilde{t}_i(t) = \infty. \tag{2.22}$$

Let $S_i(\bar{p}_k) = \tilde{S}_i(0, \bar{p}_k)$ denote the value of player i after the history (2.22).

Consider a history h_t s.t.

$$\#O_t = k, \tilde{t}_i(t) = \tilde{t}_j(t) = \infty. \tag{2.23}$$

Let $L_i(\bar{p}_k) = \tilde{L}_i(0, \bar{p}_k)$ denote the value of player i after the history (2.23) in the case when player i is the leader, and j is the follower. Let $F_i(\bar{p}_k, T_j) = \tilde{F}(0, \bar{p}_k, T_j)$ denote the value of player i after the history (2.23) in the case when j is the leader, and i is the follower.

2.6. **The Bellman equations.** We specify the six value functions defined earlier in six steps below.

Step 1. Let $\mathscr{S}_i(S_i, \bar{p}_k, t, T_i)$ be the value function of the single player

(i) had last observation at time 0;
(ii) updated her beliefs to $\bar{p}_k$ after that observation;
(iii) is still active at time t;
(iv) plans to exit at T_i unless a new information arrives earlier.

Let, as before, τ_i be the first time of the next breakdown, and let and $\mathbb{Q}_i$ be the probability measure associated with τ_i. Then

$$\begin{aligned}(2.24)\ \mathscr{S}_i(S_i,\bar{p}_k,t,T_i) &= \mathbb{E}_t^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_i}\left[\int_t^{T_i\wedge\tau_i} R_i e^{-r(s-t)}ds\right] \\ &\quad +\mathbb{E}_t^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_i}\left[\mathbb{1}_{\tau_i<T_i}e^{-r(\tau_i-t)}\left(S_i(p_{1u}(\bar{p}_k,\tau_i|\,c_{\tau_i}))-c_{\tau_i}\right)\right],\end{aligned}$$

where $\mathbb{E}_t^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_i}$ denotes the expectation conditioned on information available at time $t\geq 0$. Evidently, for any $\bar{p}_k\in[0,1]$, $\mathscr{S}_i(S_i,\bar{p}_k,0,0)=0$, $S_i(\bar{p}_k)\geq 0$, and

$$S_i(\bar{p}_k)=\sup_{T_i\geq 0}\mathscr{S}_i(S_i,\bar{p}_k,0,T_i). \tag{2.25}$$

Step 2. Let $\tilde{\mathscr{S}}_i(\pi_0,\bar{p}_0,t,T_i)$ be the value function of the single player who

(i) observed only the exit of player j at time 0;
(ii) is still active at time t;
(iii) plans to exit at T_i unless a new information arrives earlier.

Then

$$\begin{aligned}\tilde{\mathscr{S}}_i(\pi_0,\bar{p}_0,t,T_i) &= \pi_0\int_t^{T_i}R_ie^{-r(s-t)}ds+(1-\pi_0)\mathbb{E}_t^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_i}\left[\int_t^{T_i\wedge\tau_i}R_ie^{-r(s-t)}ds\right] \\ &\quad +(1-\pi_0)\mathbb{E}_t^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_i}\left[\mathbb{1}_{\tau_i<T_i}e^{-r(\tau_i-t)}\left(S_i(p_{1u}(\bar{p}_0,\tau_i|\,c_{\tau_i}))-c_{\tau_i}\right)\right],\end{aligned} \tag{2.26}$$

where S_i is defined by (2.25). Evidently, for any $(\pi_0,\bar{p}_0)\in[0,1]\times[0,1]$, $\tilde{\mathscr{S}}_i(\pi_0,\bar{p}_0,0,0)=0$, $\tilde{S}_i(\pi_0,\bar{p}_0)\geq 0$, and

$$\tilde{S}_i(\pi_0,\bar{p}_0)=\sup_{T_i\geq 0}\tilde{\mathscr{S}}_i(\pi_0,\bar{p}_0,0,T_i). \tag{2.27}$$

Notice that while problem (2.25) requires to find simultaneously the value function and the optimal exit time, problem (2.27) involves only maximization over time once one knows the value function in (2.25) for any belief $\bar{p}_k$ about the probability of the good state.

Step 3. Let $\mathscr{L}_i(L_i,\bar{p}_k,t,T_i)$ be the value function of the leader i who

(i) takes as given the pre-commitment of the follower j;
(ii) had last observation at time 0;
(iii) updated her beliefs to $\bar{p}_k$ after that observation;
(iv) is still active at time t;
(v) plans to exit at T_i unless a new information arrives earlier.

Then

$$\begin{aligned}\mathscr{L}_i(L_i,\bar{p}_k,t,T_i) &= \mathbb{E}_t^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\int_t^{T_i\wedge\tau_i\wedge\tau_j}R_ie^{-r(s-t)}ds\right] \\ &\quad +\mathbb{E}_t^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{\tau_i<T_i\wedge\tau_j}e^{-r(\tau_i-t)}\left(L_i(p_{1u}(\bar{p}_k,\tau_i|\,c_{\tau_i}))-c_{\tau_i}\right)\right] \\ &\quad +\mathbb{E}_t^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{\tau_j<T_i\wedge\tau_i}e^{-r(\tau_j-t)}L_i(p_{1u}(\bar{p}_k,\tau_j|\,c_{\tau_j}))\right].\end{aligned} \tag{2.28}$$

Evidently, for any $\bar{p}_k \in [0,1]$, $\mathscr{L}_i(L_i, \bar{p}_k, 0, 0) = 0$, $L_i(\bar{p}_k) \geq 0$, and

$$L_i(\bar{p}_k) = \sup_{T_i \geq 0} \mathscr{L}_i(L_i, \bar{p}_k, 0, T_i). \tag{2.29}$$

Step 4. Let $\tilde{\mathscr{L}}_i(\pi_0, \bar{p}_0, t, T_i)$ be the value function of the leader i who

(i) takes as given pre-commitment of the follower j;
(ii) had no observations up to time t;
(iii) is still active at time t;
(iv) plans to exit at T_i unless a new information arrives earlier.

Then

$$\begin{aligned}
\tilde{\mathscr{L}}_i(\pi_0, \bar{p}_0, t, T_i) &= \pi_0 \int_t^{T_i} R_i e^{-r(s-t)} ds + (1-\pi_0)\mathbb{E}_t^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\int_t^{T_i\wedge\tau_i\wedge\tau_j} R_i e^{-r(s-t)} ds\right] \\
&+ (1-\pi_0)\mathbb{E}_t^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{\tau_i < T_i\wedge\tau_j} e^{-r(\tau_i - t)}\left(L_i(p_{1u}(\bar{p}_0, \tau_i | \, c_{\tau_i})) - c_{\tau_i}\right)\right] \\
&+ (1-\pi_0)\mathbb{E}_t^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{\tau_j < T_i\wedge\tau_i} e^{-r(\tau_j - t)} L_i(p_{1u}(\bar{p}_0, \tau_j | \, c_{\tau_j}))\right].
\end{aligned} \tag{2.30}$$

Evidently, for any $(\pi_0, \bar{p}_0) \in [0,1]\times[0,1]$, $\tilde{\mathscr{L}}_i(\pi_0, \bar{p}_0, 0, 0) = 0$, $\tilde{L}_i(\pi_0, \bar{p}_0) \geq 0$, and

$$\tilde{L}_i(\pi_0, \bar{p}_0) = \sup_{T_i \geq 0} \tilde{\mathscr{L}}_i(\pi_0, \bar{p}_0, 0, T_i). \tag{2.31}$$

Notice that once we solve problem (2.29) and obtain the value function $L_i(\bar{p}_k)$, problem (2.31) becomes a maximization problem w.r.t. T_i.

Step 5. Let us consider the value function of the follower i who

(i) takes as given the time of exit T_j of the leader j;
(ii) had last observation at time 0;
(iii) updated her beliefs to $\bar{p}_k$ after that observation;
(iv) plans to exit at $T_i \geq T_j$ unless a new information arrives before T_i.

Then

$$\begin{aligned}
F_i(\bar{p}_k, T_j) &= \mathbb{E}_0^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\int_0^{T_j\wedge\tau_i\wedge\tau_j} R_i e^{-rs} ds\right] \\
&+\mathbb{E}_0^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{\tau_i < T_j\wedge\tau_j} e^{-r\tau_i}\left(F_i(p_{1u}(\bar{p}_k, \tau_i | \, c_{\tau_i}), T_j) - c_{\tau_i}\right)\right] \\
&+\mathbb{E}_0^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{\tau_j < T_j\wedge\tau_i} e^{-r\tau_j} F_i(p_{1u}(\bar{p}_k, \tau_j | \, c_{\tau_j}), T_j)\right] \\
&+\mathbb{E}_0^{\mathbb{P}(\bar{p}_k)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2}\left[\mathbb{1}_{T_j < \tau_j\wedge\tau_i} e^{-rT_j} S_i(p_1(\bar{p}_k, T_j))\right].
\end{aligned} \tag{2.32}$$

Notice that problem (2.32) does not contain maximization over time. One only needs to find the unknown value function $F_i(\bar{p}_k, T_j)$.

Step 6. Finally, let us consider the value function of the follower i who

(i) takes as given the time of exit T_j of the leader j;
(ii) had no observations;
(iii) plans to exit at $T_i \geq T_j$ unless a new information arrives before T_i.

Then

$$
\begin{aligned}
\tilde{F}_i(\pi_0, \bar{p}_0, T_j) \;=\;& \pi_0 \left(\int_0^{T_j} R_i e^{-rs} ds + e^{-rT_j} \tilde{S}_i(\pi_0, p_1(\bar{p}_0, T_j)) \right) \\
&+(1-\pi_0)\mathbb{E}_0^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2} \left[\int_0^{T_j\wedge\tau_i\wedge\tau_j} R_i e^{-rs} ds \right] \\
&+(1-\pi_0)\mathbb{E}_0^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2} \left[\mathbb{1}_{\tau_i<T_j\wedge\tau_j} e^{-r\tau_i} \left(F_i(p_{1u}(\bar{p}_0, \tau_i | \, c_{\tau_i}), T_j) - c_{\tau_i} \right) \right] \\
&+(1-\pi_0)\mathbb{E}_0^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2} \left[\mathbb{1}_{\tau_j<T_j\wedge\tau_i} e^{-r\tau_j} F_i(p_{1u}(\bar{p}_k, \tau_j | \, c_{\tau_j}), T_j) \right] \\
&+(1-\pi_0)\mathbb{E}_0^{\mathbb{P}(\bar{p}_0)\otimes\mathbb{Q}_1\otimes\mathbb{Q}_2} \left[\mathbb{1}_{T_j<\tau_j\wedge\tau_i} e^{-rT_j} S_i(p_1(\bar{p}_0, T_j)) \right].
\end{aligned}
\tag{2.33}
$$

Notice that the RHS in (2.33) contains the value functions, which were defined in the previous steps, hence no additional optimization is required.

3. Main steps of solution

In this Section, we explain how to find the value functions and optimal stopping strategies defined in the previous Section.

3.1. **Value of the single player after at least one breakdown.** For the time being, we will suppress the subscript i. Let $\bar{p}_k$ be the updated posterior at the moment of the last breakdown. Denote by $U = \{\bar{p}_k \in [0,1] | \, T^{\bar{p}_k} > 0\}$ the waiting region of the player. We start with the following result.

Lemma 3.1. *(i) If $\exists \, \bar{p}_k \in [0,1]$ s.t. $S(\bar{p}_k) = 0$, the waiting region is of the form $U = (\bar{p}^*, 1]$, where $\bar{p}^*$ is the supremum of the set of all such $\bar{p}_k$. The value function $S(\bar{p}_k)$ is increasing on U. (ii) If $S(\bar{p}_k) > 0$ for any $\bar{p}_k \in [0,1]$, it is never optimal to exit until a new observation.*

Proof. (i) Consider $\bar{p}_k < \bar{p}'_k$ s.t. $S(\bar{p}_k) > 0$. Let $T^{\bar{p}_k} > 0$ be the optimal exit time in a subgame that starts at $(0, \bar{p}_k)$. Suppose that, in a subgame that starts at $(0, \bar{p}'_k)$, the player wants to exit at time $T^{\bar{p}_k}$, which may be suboptimal. With a positive probability the next cost will be incurred at random time $\tau < T^{\bar{p}_k}$. Let $\mathbb{E}[c_i]$ denote the expected cost if the nature is in state t at the moment of observation. The expected cost at τ is

$$
\begin{aligned}
p_1(\bar{p}'_k, \tau)\mathbb{E}[c_1] + p_2(\bar{p}'_k, \tau)\mathbb{E}[c_2] &= p_1(\bar{p}'_k, \tau)(\mathbb{E}[c_1] - \mathbb{E}[c_2]) + \mathbb{E}[c_2] \\
&< p_1(\bar{p}_k, \tau)(\mathbb{E}[c_1] - \mathbb{E}[c_2]) + \mathbb{E}[c_2],
\end{aligned}
$$

because

$$
\begin{aligned}
\mathbb{E}[c_1] - \mathbb{E}[c_2] &= (1-\alpha_1)c_l + (\alpha_1 - \alpha_2)c_m(1-\alpha_2)c_h \\
&= \alpha_1(c_m - c_l) + \alpha_2(c_h - c_m) - c_h \\
&< \min\{\alpha_1, \alpha_2\}(c_h - c_l) - (c_h - c_l) < 0,
\end{aligned}
$$

and $p_1(\bar{p}'_k, \tau) > p_1(\bar{p}_k, \tau)$ on the strength of Claim 2.1. Hence the value function at $\bar{p}'_k$ with the suboptimal exit rule $T^{\bar{p}_k}$ is greater than $S(\bar{p}_k)$. Choosing the optimal exit time $T^{\bar{p}'_k}$ cannot decrease the value. Therefore $S(\bar{p}'_k) > S(\bar{p}_k) > 0$, which implies

that S is increasing on $(\bar{p}^*, 1]$. The same argument shows that if $S(\bar{p}_k) > 0$ for some $\bar{p}_k < \bar{p}^*$, then $S(\bar{p}_k^*) > 0$, which is a contradiction. Hence $S(\bar{p}_k) = 0$ for all $\bar{p}_k \leq \bar{p}^*$. (ii) If $S(\bar{p}_k) > 0$ for any $\bar{p}_k \in [0, 1]$, then we set $\bar{p}^* = 0$, $U = [0, 1]$, and $T^{\bar{p}_k} = \infty$ for any $\bar{p}_k \in U$, i.e., it is never optimal to exit until a new observation arrives.

□

Let us consider the optimization problem (2.25) in more detail. Standard considerations show that the value function $\mathscr{S}$ defined by (2.24) can be written as

$$\mathscr{S}(S, \bar{p}_k, t, T) = \int_t^T e^{-(r+\lambda)(s-t)} \left(R + \lambda \mathbb{E}\left[S(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right]\right) ds, \tag{3.1}$$

where

$$\begin{aligned} \mathbb{E}\left(S(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right) &= (1-\alpha_1)p_1(\bar{p}_k, s)(S(1) - c_l) \\ &+ (1-\alpha_2)p_2(\bar{p}_k, s)(S(0) - c_h) \\ &+ (\alpha_1 p_1(\bar{p}_k, s) + \alpha_2 p_2(\bar{p}_k, s))(S(p_{1u}(\bar{p}_k, s|\, c_s = c_m)) - c_m). \end{aligned} \tag{3.2}$$

Set $S_0(\bar{p}_k) = 0$, and define, inductively, for $n = 0, 1, 2, \ldots$,

$$S_{n+1}(\bar{p}_k) = \sup_{T \geq 0} \mathscr{S}(S_n, \bar{p}_k, 0, T), \quad \bar{p}_k \in [0, 1]. \tag{3.3}$$

Theorem 3.2. *The following statements hold:*

(i) The sequence $\{S_n\}_{n\in\mathbb{N}}$ is non-decreasing (point-wise).
(ii) The sequence $\{S_n\}_{n\in\mathbb{N}}$ is uniformly bounded from above by R/r.
(iii) For any $\bar{p}_k \in [0, 1]$, the limit $S(\bar{p}_k) = \lim_{n\to\infty} S_n(\bar{p}_k)$ exists, S admits the bound $0 \leq S(\bar{p}_k) \leq R/r$, (2.25) holds, and S is the value function.

Proof. (i) is evident from (3.1) (ii) It suffices to prove that $\mathscr{S}(S, \bar{p}_k, t, T)$ is bounded by R/r if S on the RHS of (3.1) is bounded by R/r. Replacing S with R/r and omitting the expected cost in (3.1), we obtain

$$\begin{aligned} \mathscr{S}(S, \bar{p}_k, t, T) &\leq \int_t^T e^{-(r+\lambda)(s-t)} ds \left[R + \lambda \frac{R}{r}\right] \\ &= \frac{R(r+\lambda)}{r} \cdot \frac{1 - e^{-(r+\lambda)(T-t)}}{r+\lambda} \leq \frac{R}{r}. \end{aligned}$$

(iii) The existence of the limit and bounds on S follow from (i) and (ii). Passing to the limit in (3.13), we obtain (2.25). Hence, S is the value function. □

In order to study the set of the optimal $T = T^{\bar{p}_k}$, set $f(t) = e^{-(r+\lambda)t}$,

$$\mathcal{U}(S, \bar{p}_k, s) = R + \lambda \mathbb{E}\left[S(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right], \tag{3.4}$$

then (3.1) becomes

$$\mathscr{S}(S, \bar{p}_k, t, T) = \int_t^T \frac{f(s)}{f(t)} \mathcal{U}(S, \bar{p}_k, s) ds. \tag{3.5}$$

Calculate the derivative $\mathscr{S}_T := \partial \mathscr{S}(S, \bar{p}_k, t, T)/\partial T$:

$$\mathscr{S}_T(S, \bar{p}_k, t, T) = \frac{f(T)}{f(t)} \mathcal{U}(S, \bar{p}_k, T).$$

Clearly, $\mathscr{S}_T(S, \bar{p}_k, t, T) > 0 (= 0, < 0) \Leftrightarrow \mathcal{U}(S, \bar{p}_k, T) > 0 (= 0, < 0)$. Notice that $\mathcal{U}(S, \bar{p}_k, T)$ is the difference between the marginal benefit $R + \lambda \mathbb{E}\left[S(p_{1u}(\bar{p}_k, T | \, c_T))\right]$ and the marginal cost $\lambda \mathbb{E}[c_T]$ of being active at time T. If $T^{\bar{p}_k}$ is an interior maximum of $\mathscr{S}(S, \bar{p}_k, t, T)$, then $T^{\bar{p}_k}$ satisfies $\mathcal{U}(S, \bar{p}_k, T) = 0$, i.e., it is optimal to exit when the expected marginal benefit equals to the expected marginal cost.

Theorem 3.3. *An optimal $T = T^{\bar{p}_k}$ is time-consistent: if $T^{\bar{p}_k}$ maximizes $\mathscr{S}(S, \bar{p}_k, 0, T)$, then, for any $t \le T^{\bar{p}_k}$, $T^{\bar{p}_k}$ is a maximizer of $\mathscr{S}(S, \bar{p}_k, t, T)$.*

Proof. For any $0 \le t < T$, we have

$$\begin{aligned}
\mathscr{S}(S, \bar{p}_k, 0, T) &= \int_0^T \frac{f(s)}{f(0)} \mathcal{U}(S, \bar{p}_k, s) ds \\
&= \int_0^t \frac{f(s)}{f(0)} \mathcal{U}(S, \bar{p}_k, s) ds + \frac{f(t)}{f(0)} \int_t^T \frac{f(s)}{f(0)} \mathcal{U}(S, \bar{p}_k, s) ds \\
&= \mathscr{S}(S, \bar{p}_k, 0, t) + \frac{f(t)}{f(0)} \mathscr{S}(S, \bar{p}_k, t, T).
\end{aligned}$$

Since $\mathscr{S}(S, \bar{p}_k, 0, t)$ and $f(t)/f(0) (> 0)$ are independent of T, a T which maximizes $\mathscr{S}(S, \bar{p}_k, 0, T)$ maximizes $\mathscr{S}(S, \bar{p}_k, t, T)$ as well. □

Theorem 3.4. *Let the following conditions hold*

$$\lambda \mathbb{E}[c_1] < R < \frac{r\lambda}{r + \lambda} \mathbb{E}[c_2], \tag{3.6}$$

$$\frac{\mu_2}{\mu_1 + \mu_2} < \frac{\lambda \mathbb{E}[c_2] - \frac{r+\lambda}{r} R}{\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1])}. \tag{3.7}$$

Then

(i) $\exists \, T = T^{\bar{p}_k} > 0$ s.t.

$$\mathcal{U}(S, \bar{p}_k, T^{\bar{p}_k}) = 0; \tag{3.8}$$

(ii) an optimal exit time $T_*^{\bar{p}_k}$ is a solution of (3.8);

(iii) the cutoff belief (the left boundary of the waiting region U) satisfies

$$\bar{p}^* < \frac{\lambda \mathbb{E}[c_2] - R}{\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1])}. \tag{3.9}$$

Conditions specified in Theorem 3.4 admit the following interpretation. The first inequality in (3.16) says that the instantaneous revenue is higher than the instantaneous expected cost in the "good" state. The LHS in inequality (3.17) is the long-run belief about the probability of "good" state, $\lim_{t \to \infty} p_1(\bar{p}_k, t)$. Hence if the long-run belief about the probability of "good" state is sufficiently low, the player will exit in finite time. The second inequality in (3.16), which says that the present value of the

perpetual revenue flow R/r is less than the expected value of life-time losses in the "bad" state, ensures that the set of long-run beliefs that satisfy (3.17) is non-empty.

Proof. It is straightforward to show that $\mathcal{U}$ is continuous in T. To prove that there exists a belief $\bar{p}_k$, s.t. $T_*^{\bar{p}_k} > 0$, we show that for any $\bar{p}_k$ s.t.

$$\bar{p}_k \geq \frac{\lambda\mathbb{E}[c_2] - R}{\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1])} \tag{3.10}$$

$\mathcal{U}(S, \bar{p}_k, +0) > 0$. Then $\mathscr{S}$ is increasing in T in a neighborhood of 0, and immediate exit is not optimal. It is easy to see that the first inquality in (3.16) ensures that the set of beliefs $\bar{p}_k$ that satisfy (3.10) is non-empty.

$$\begin{aligned}\mathcal{U}(S, \bar{p}_k, +0) &= R - \lambda\mathbb{E}[c_2] + \lambda\bar{p}_k(\mathbb{E}[c_2] - \mathbb{E}[c_1]) + \lambda\mathbb{E}[S(p_1(\bar{p}_k, +0))] \\ &\geq R - \lambda\mathbb{E}[c_2] + \lambda\bar{p}_k(\mathbb{E}[c_2] - \mathbb{E}[c_1]).\end{aligned}$$

Hence, if (3.16) holds as a strict inequality, $\mathcal{U}(S, \bar{p}_k, +0) > 0$. Consider

$$\bar{p}'_k = \frac{\lambda\mathbb{E}[c_2] - R}{\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1])}$$

and calculate $\mathcal{U}(S, \bar{p}'_1, \Delta t)$. Since

$$p_1(\bar{p}'_k, \Delta t) = \bar{p}'_k - (\bar{p}'_k(\mu_1 + \mu_2) - \mu_2)\Delta t + o(\Delta t),$$

we can write

$$\begin{aligned}\mathcal{U}(S, \bar{p}'_k, \Delta t) &= \lambda\left[\bar{p}'_k(1-\alpha_1)S(1) + (1-\bar{p}'_k)(1-\alpha_2)S(0)\right. \\ &\quad \left. +(\alpha_2 + (\alpha_1 - \alpha_2)\bar{p}'_k)S\left(\frac{\alpha_1\bar{p}'_k}{\alpha_2 + (\alpha_1 - \alpha_2)\bar{p}'_k}\right)\right] \\ &\quad -\lambda(\bar{p}'_k(\mu_1 + \mu_2) - \mu_2)\Delta t\left[\mathbb{E}[c_2] - \mathbb{E}[c_1] + (1-\alpha_1)S(1) - (1-\alpha_2)S(0)\right] \\ &\quad -\lambda(\bar{p}'_k(\mu_1 + \mu_2) - \mu_2)\Delta t S\left(\frac{\alpha_1\bar{p}'_k}{\alpha_2 + (\alpha_1 - \alpha_2)\bar{p}'_k}\right) + o(\Delta t).\end{aligned}$$

Passing to the limit at $\Delta t \downarrow 0$, we obtain

$$\begin{aligned}\mathcal{U}(S, \bar{p}'_k, +0) &= \lambda\left[\bar{p}'_k(1-\alpha_1)S(1) + (1-\bar{p}'_k)(1-\alpha_2)S(0)\right. \\ &\quad \left. +(\alpha_2 + (\alpha_1 - \alpha_2)\bar{p}'_k)S\left(\frac{\alpha_1\bar{p}'_1}{\alpha_2 + (\alpha_1 - \alpha_2)\bar{p}'_k}\right)\right] \geq \lambda\bar{p}'_k(1-\alpha_1)S(1) > 0.\end{aligned}$$

Therefore the cutoff belief satisfies (3.19), and (iii) follows.

Since $S(\bar{p}_k) \leq R/r$, we can write that

$$\begin{aligned}\lim_{T\to\infty}\mathcal{U}(S, \bar{p}_k, T) &= R - \lambda\lim_{T\to\infty}\mathbb{E}[c_T] + \lambda\lim_{T\to\infty}\mathbb{E}[S(p_{1,up}(\bar{p}_k, T|\, c_T = c_m)] \\ &\leq \frac{r+\lambda}{r}R - \lambda\mathbb{E}[c_2] + \frac{\mu_2}{\mu_1 + \mu_2}\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1]) < 0\end{aligned}$$

on the strength of (3.17). Hence, $\mathcal{U}$ is negative in a neighborhood of $+\infty$. Since $\mathcal{U}$ is positive in a neighborhood of 0, all points of global maximum are interior, and the FOC (3.18) is satisfied at these points, which proves (i). □

3.2. **Value of the leader after at least one breakdown.** We will focus first on the case of heterogeneous agents, assuming that the revenue stream $R_2 > R_1$. The next result shows that the players' roles are predetermined in this case.

Let $\Pi_i(\bar{p}_k, t)$ denote the expected profit of player i when this player is active at time t after the last update $\bar{p}_k$. Then, for every i and $t > 0$,

$$\begin{aligned} \Pi_1(\bar{p}_k, t) &= \int_0^t e^{-rs} \left(R_1 - \lambda \mathbb{E}[c_s]\right) ds \\ &< \int_0^t e^{-rs} \left(R_2 - \lambda \mathbb{E}[c_s]\right) ds = \Pi_2(\bar{p}_k, t). \end{aligned} \tag{3.11}$$

Lemma 3.5. *Let $\Pi_1(\bar{p}_k, t) \le \Pi_2(\bar{p}_k, t)$ $\forall\, (\bar{p}_k, t)$ and $T_1^{\bar{p}_k} > 0$. Then $T_1^{\bar{p}_k} \le T_2^{\bar{p}_k}$ $\forall\, \bar{p}_k$.*

Proof. Let τ^{12} denote the random time when a new breakdown happens to one of the players. Consider $0 < t < T_1^{\bar{p}_k} \wedge \tau^{12}$. Then

$$\mathbb{E}_{0,\bar{p}_k}\left[e^{-rt} V_1(0, \bar{p}_k)\right] > 0.$$

Since $\Pi_1(\bar{p}_k, t) \le \Pi_2(\bar{p}_k, t)$, $\mathbb{E}_{0,\bar{p}_k}\left[e^{-rt} V_2(0, \bar{p}_k)\right] \ge \mathbb{E}_{0,\bar{p}_k}\left[e^{-rt} V_1(0, \bar{p}_k)\right]$, hence exit at t cannot be optimal for player 2, hence $T_1^{\bar{p}_k} \le T_2^{\bar{p}_k}$. ◻

Since the roles of the players are predetermined by the assumptions of the model, we see that both players will make non-strategic exit decisions.

Let us consider the optimization problem (2.29) in more detail. Standard considerations show that the value function $\mathscr{L}_1$ defined by (2.28) can be written as

$$\mathscr{L}_1(L_1, \bar{p}_k, t, T_1) = \int_t^{T_1} e^{-(r+2\lambda)(s-t)} \left(R_1 + \lambda \mathbb{E}\left[2L_1(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right]\right) ds. \tag{3.12}$$

Set $L_{1,0}(\bar{p}_k) = 0$, and define, inductively, for $n = 0, 1, 2, \dots$,

$$L_{1,n+1}(\bar{p}_k) = \sup_{0 \le T_1} \mathscr{L}_1(L_{1,n}, \bar{p}_k, 0, T_1), \quad \bar{p}_k \in [0,1]. \tag{3.13}$$

Theorem 3.6. *The following statements hold:*

(i) The sequence $\{L_{1,n}\}_{n\in\mathbb{N}}$ is non-decreasing (point-wise).
(ii) The sequence $\{L_{1,n}\}_{n\in\mathbb{N}}$ is uniformly bounded from above by R/r.
(iii) For any $\bar{p}_k \in [0,1]$, the limit $L_1(\bar{p}_k) = \lim_{n\to\infty} L_{1,n}(\bar{p}_k)$ exists, L_1 admits the bound $0 \le L_1(\bar{p}_k) \le R/r$, (2.29) holds, and L_1 is the value function.

The proof repeats word by word the proof of Theorem 3.2 if one replaces S in that proof with L_1.

One can also repeat the statement and proof of Lemma 3.1, replacing S with L_1 to establish the fact that the waiting region is of the form $U_d = (\bar{p}_d^*, 1]$.

In order to study the set of the optimal $T_1 = T_1^{\bar{p}_k}$, set $f(t) = e^{-(r+2\lambda)t}$,

$$\mathscr{U}_1(L_1, \bar{p}_k, s) = R_1 + \lambda \mathbb{E}\left[2L_1(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right], \tag{3.14}$$

then (3.12) becomes

$$\mathscr{L}_1(L, \bar{p}_k, t, T_1) = \int_t^{T_1} \frac{f(s)}{f(t)} \mathcal{U}_1(L_1, \bar{p}_k, s) ds. \tag{3.15}$$

Calculate the derivative

$$\frac{\partial \mathscr{L}_1(L_1, \bar{p}_k, t, T_1)}{\partial T_1} = \frac{f(T_1)}{f(t)} \mathcal{U}_1(L_1, \bar{p}_k, T_1).$$

If $T_1^{\bar{p}_k}$ is an interior maximum of $\mathscr{L}_1(L_1, \bar{p}_k, t, T_1)$, then $T_1^{\bar{p}_k}$ satisfies $\mathcal{U}_1(L_1, \bar{p}_k, T_1) = 0$, i.e., it is optimal to exit when the expected marginal benefit equals to the expected marginal cost.

Theorem 3.7. *An optimal $T_1 = T_1^{\bar{p}_k}$ is time-consistent: if $T_1^{\bar{p}_k}$ maximizes $\mathscr{L}_1(L_1, \bar{p}_k, 0, T_1)$, then, for any $t \le T_1^{\bar{p}_k}$, $T_1^{\bar{p}_k}$ is a maximizer of $\mathscr{L}_1(L_1, \bar{p}_k, t, T_1)$*

The proof repeats the proof of Theorem 3.3.

Theorem 3.8. *Let the following conditions hold*

$$\lambda \mathbb{E}[c_1] < R_1 < \frac{r\lambda}{r + 2\lambda} \mathbb{E}[c_2], \tag{3.16}$$

$$\frac{\mu_2}{\mu_1 + \mu_2} < \frac{\lambda \mathbb{E}[c_2] - \frac{r+2\lambda}{r} R_1}{\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1])}. \tag{3.17}$$

Then

(i) $\exists\, T_1 = T_1^{\bar{p}_k} > 0$ s.t.

$$\mathcal{U}_1(L_1, \bar{p}_k, T_1^{\bar{p}_k}) = 0; \tag{3.18}$$

(ii) an optimal exit time $T_{1*}^{\bar{p}_k}$ is a solution of (3.18)*;*
(iii) the cutoff belief (the left boundary of the waiting region U_d) satisfies

$$\bar{p}_d^* < \frac{\lambda \mathbb{E}[c_2] - R}{\lambda(\mathbb{E}[c_2] - \mathbb{E}[c_1])}. \tag{3.19}$$

The proof repeats the proof of Theorem 3.4.

It is interesting to compare the waiting regions of player 1, when she is a single player, and when she is the leader. We were unable to prove it, but observed in numerical experiments that $L_1(\bar{p}_k) \ge S_1(\bar{p}_k)$ for all $\bar{p}_k$. Provided this conjecture is correct, it is easy to show that $\bar{p}_d^* < \bar{p}^*$. Indeed,

$$\begin{aligned}
\mathcal{U}(S, \bar{p}^*, +0) &= R_1 - \lambda \mathbb{E}[c_2] + \lambda \bar{p}^*(\mathbb{E}[c_2] - \mathbb{E}[c_1]) + \lambda \mathbb{E}[S_1(p_1(\bar{p}^*, +0))] \\
&< R_1 - \lambda \mathbb{E}[c_2] + \lambda \bar{p}^*(\mathbb{E}[c_2] - \mathbb{E}[c_1]) + 2\lambda \mathbb{E}[S_1(p_1(\bar{p}^*, +0))] \\
&\le R_1 - \lambda \mathbb{E}[c_2] + \lambda \bar{p}^*(\mathbb{E}[c_2] - \mathbb{E}[c_1]) + 2\lambda \mathbb{E}[L_1(p_1(\bar{p}^*, +0))] \\
&= \mathcal{U}_1(L_1, \bar{p}^*, +0).
\end{aligned}$$

It follows that $\mathcal{U}_1(L_1, \bar{p}^*, +0) > 0$, hence $\bar{p}_d^* < \bar{p}^*$. See Fig. 1 for illustration. This result differs from Heidhues et al. [21], where all the players stop experimenting at the individually optimal belief threshold if information about payoffs is public. This

result in [21] is due to a severe free riding problem, which does not exist in our model.

3.3. **Value of the follower after at least one breakdown.**

The player's payoff asymmetry implies that player 2 is the follower. Let us consider problem (2.32) in more detail. Standard considerations show that the value function $F_2(\bar{p}_k, T_1)$ in (2.32) can be written as

$$\begin{aligned} F_2(\bar{p}_k, T_1) &= \int_0^{T_1} e^{-(r+2\lambda)(s-t)} \left(R_2 + \lambda \mathbb{E}\left[2F_2(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right]\right) ds \\ &\quad + e^{-(r+2\lambda)T_1} S_2(p_1(\bar{p}_k, T_1)). \end{aligned} \tag{3.20}$$

Set $F_{2,0}(\bar{p}_k, T_1) = 0$, and define, inductively, for $n = 0, 1, 2, \ldots$,

$$\begin{aligned} F_{2,n}(\bar{p}_k, T_1) &= \int_0^{T_1} e^{-(r+2\lambda)(s-t)} \left(R_2 + \lambda \mathbb{E}\left[2F_{2,n-1}(p_{1u}(\bar{p}_k, s|\, c_s)) - c_s\right]\right) ds \\ &\quad + e^{-(r+2\lambda)T_1} S_2(p_1(\bar{p}_k, T_1)). \end{aligned} \tag{3.21}$$

Theorem 3.9. *The following statements hold:*

(i) The sequence $\{F_{2,n}\}_{n\in\mathbb{N}}$ is non-decreasing (point-wise).
(ii) The sequence $\{F_{2,n}\}_{n\in\mathbb{N}}$ is uniformly bounded from above by R/r.
(iii) For any $\bar{p}_k \in [0,1]$, the limit $F_1(\bar{p}_k, T_1, T_2) = \lim_{n\to\infty} F_{1,n}(\bar{p}_k, T_1, T_2)$ exists, L_2 admits the bound $0 \le L_1(\bar{p}_k) \le R/r$, (2.32) holds, and L_1 is the value function.

The proof repeats word by word the proof of Theorem 3.2 if one replaces S in that proof with F_2. See Fig. 2 for illustration.

4. Smooth or not smooth

In this section, we study the regularity of the value function of a single player at the boundary of the stopping region.

4.1. **Value after at least one breakdown.**

Recall that the stopping region is of the form $U^c = [0, \bar{p}^*]$.

Theorem 4.1. *Let $\bar{p}^* > \frac{\mu_2}{\mu_1+\mu_2}$. Then*
a) for any $\bar{p}_k$, the optimal exit time $T^*(\bar{p}_k) < \infty$;
b) S is smooth at $\bar{p}^*$.

Proof. First, we prove that the value function S is Lipschitz continuous at $\bar{p}^*$. Let $\bar{p}_k > \bar{p}^*$. Since $\bar{p}^*$ is higher than the long run probability $\frac{\mu_2}{\mu_1+\mu_2}$, the right derivative of $t \mapsto p_1(\bar{p}_k, t)$ at 0 is negative uniformly in $\bar{p}_k \in [\bar{p}^*, 1]$. This implies a).

To prove b), for $\bar{p}_k > \bar{p}^*$, set $\epsilon = \bar{p}^* - \bar{p}^*$, and note that there exists $a > 0$ independent of ϵ such that $T^*(\bar{p}_k) \le a\epsilon$. Let τ be the first cost observation. The probability of the event $\{\tau \le T^*(\bar{p}_k)\}$ is bounded by $b\epsilon$, where b is independent of ϵ, and the cost c_τ and the value function at time τ are uniformly bounded. Hence,

$$0 < S(\bar{p}_k) - S(\bar{p}^*) = S(\bar{p}_k) \le (1 - e^{-rT^*(\bar{p}_k)})R/r + A_1\epsilon \le A_2(\bar{p}_k - \bar{p}^*), \tag{4.1}$$

where $A_1, A_2 > 0$ are independent of $\bar{p}_k$. Similarly, the present value of costs is $O(\bar{p}_k - \bar{p}^*)$, and we conclude that, uniformly in $t \in (0, T^*(\bar{p}_k))$,

$$\mathcal{U}(S, \bar{p}_k, t) = \mathcal{U}(S, \bar{p}_k, t) - \mathcal{U}(S, \bar{p}_k, T^*(\bar{p}_k)) = O(\bar{p}_k - \bar{p}^*).$$

Since the exit time $T^*(\bar{p}_k) \downarrow 0$ as $\bar{p}_k \downarrow \bar{p}^*$,

$$S(\bar{p}_k) = \int_0^{T^*(\bar{p}_k)} e^{-(r+2\lambda)t} \mathcal{U}(S, \bar{p}_k, t) dt = o(\bar{p}_k - \bar{p}^*).$$

□

Theorem 4.2. *Let $\bar{p}^*$ be the cutoff belief, and $\bar{p}^* < \frac{\mu_2}{\mu_1+\mu_2}$. Then*

a) for any $\bar{p}_k > \bar{p}^*$, the optimal exit time $T^*(\bar{p}_k) = \infty$.

b) S is non-smooth at $\bar{p}^*$. Specifically, there exists $a > 0$ such that, for all $\bar{p}_k > \bar{p}^*$,

$$S(\bar{p}_k) \geq a(\bar{p}_k - \bar{p}^*). \tag{4.2}$$

Proof. a) If $\bar{p}_k > \frac{\mu_2}{\mu_1+\mu_2}$, the deterministic belief $p_1(\bar{p}_k, t) > \frac{\mu_2}{\mu_1+\mu_2} > \bar{p}^*$ for all t, hence $S(p_1(\bar{p}_k, t)) > 0$ for all t, and it is non-optimal to exit earlier than a cost is observed. If $\bar{p}_k \in (\bar{p}^*, \frac{\mu_2}{\mu_1+\mu_2})$, $t \mapsto p_1(\bar{p}_k, t)$ increases, hence, $S(p_1(\bar{p}_k, t)) > 0$ for all t, and it is non-optimal to exit earlier than a cost is observed.

b) Consider two sample paths of beliefs $p_1(\bar{p}_k, \omega, t)$ and $p_1(\bar{p}'_1, \omega, t)$, where $\bar{p}_k > \bar{p}'_k > \bar{p}^*$. After the first conclusive observation τ, the sample paths $p_1(\bar{p}_k, \omega, t)$ and $p_1(\bar{p}'_1, \omega, t)$ coincide, and, for $t \in (0, \tau)$,

$$p_1(\bar{p}_k, \omega, t) - p_1(\bar{p}'_k, \omega, t) > a(\bar{p}_k - \bar{p}'_k), \tag{4.3}$$

where $a > 0$ is independent of $\bar{p}_k > \bar{p}'_k > \bar{p}^*$ and the time of the first conclusive observation. (This follows from the formulas for the evolution of the deterministic belief before the first conclusive observation.)

The probability that the first observation τ_1 is inconclusive is positive, and, on the strength of (4.3),

$$\mathbb{E}_{\bar{p}'_k;\tau_1}[c_{\tau_1}] - \mathbb{E}_{\bar{p}_k;\tau_1}[c_{\tau_1}] \geq b(\bar{p}_k - \bar{p}'_k),$$

where $a > 0$ is independent of $\bar{p}_k > \bar{p}'_k > \bar{p}^*$ and the time of the first inconclusive observation. Hence, there exists $d > 0$ such that, for all $\bar{p}_k > \bar{p}'_k > \bar{p}^*$, $S(\bar{p}_k) - S(\bar{p}'_k) \geq d(\bar{p}_k - \bar{p}'_k)$. Passing to the limit $\bar{p}'_k \downarrow \bar{p}^*$, we obtain (4.2). □

Fig. 3 illustrates both theorems. We show two graphs of the value functions (the smooth pasting and kink) but only one graph for T^* because in the second case $T^*(\bar{p}_k) = 0$ for $\bar{p}_k \leq \bar{p}^*$, and $T^*(\bar{p}_k) = \infty$ for $\bar{p}_k > \bar{p}^*$.

4.2. **Value before breakdown.** In this case, the stopping region U^c is a subset of $[0, 1] \times [0, 1]$. We formulate analogs of Theorems 4.1 and 4.2 in terms of sample paths of deterministic beliefs and parts of the boundary ∂U^c of the stopping region.

Definition 4.3. A subset $\partial U^c_{\text{absorb}} \subset \partial U^c$ is called *absorbing* if there exists a neighborhood of $\partial U^c_{\text{absorb}}$ in U such that each deterministic trajectory that starts in the neighborhood enters U^c.

A subset $\partial U^c_{\text{reflect}} \subset \partial U^c$ is called *reflecting* if, for any $(\pi^*, \bar{p}^*) \in \partial U^c_{\text{reflect}}$ there exist $\epsilon > 0$ and a neighborhood of $(\pi^*, \bar{p}^*)$ in U such that each deterministic trajectory that starts in the neighborhood does not enter U^c or enters not earlier than at $t = \epsilon$.

For $(\pi^*, \bar{p}^*) \in \partial U^c$, define the vector of derivatives of the deterministic beliefs $v(\pi^*, \bar{p}^*) := (\partial_t \pi(\pi^*, 0), \partial_t p_1(\bar{p}^*, 0))$. If ∂U^c is smooth at $(\pi^*, \bar{p}^*)$, define $n(\pi^*, \bar{p}^*)$ as the normal vector to ∂U^c pointing into U.

Theorem 4.4. *Assume that $(\pi^*, \bar{p}^*) \in \partial U^c$, and $\nu(\pi^*, \bar{p}^*) \neq 0$. Then*

a) If ∂U^c is smooth in a neighborhood of $(\pi^*, \bar{p}^*)$ and the angle between $n(\pi^*, \bar{p}^*)$ and $\nu(\pi^*, \bar{p}^*)$ is obtuse, then

- $(\pi^*, \bar{p}^*) \in \partial U^c_{\text{absorb}}$*;*
- *along any trajectory of the deterministic beliefs that starts in a sufficiently small neighborhood of $(\pi^*, \bar{p}^*)$ in U, the value function is smooth;*
- *for any $(\pi, \bar{p}_k)$ in a sufficiently small neighborhood of $(\pi^*, \bar{p}^*)$ in U, $T^*(\pi, \bar{p}_k) < \infty$.*

b) If ∂U^c is smooth in a neighborhood of $(\pi^*, \bar{p}^*)$ and the angle between $n(\pi^*, \bar{p}^*)$ and $\nu(\pi^*, \bar{p}^*)$ is sharp, then

- $(\pi^*, \bar{p}^*) \in \partial U^c_{\text{reflect}}$*;*
- *along any trajectory that a sufficiently small neighborhood of $(\pi^*, \bar{p}^*)$ in ∂U, the value function is non-smooth;*
- *for any $\epsilon > 0$, there exists a neighborhood of $(\pi^*, \bar{p}^*)$ in U^c such that, for all $(\pi, \bar{p}_k)$ in the neighborhood, $T^*(\pi, \bar{p}_k) > \epsilon$.*

The proofs of the parts a) and b) are straightforward modifications of the proofs of Theorems 4.1 and 4.2.

Remark 4.5. a) The part b) can be easily generalized to the case when ∂U^c has kinks.
b) We cannot exclude cases (for more general dynamics of beliefs, at least) when a trajectory of deterministic beliefs that starts at the reflecting boundary eventually enters the stopping region. For this trajectory, $T^* < \infty$.

See Fig. 4, where we show the trajectories that enter the stopping region and the ones that never enter, including the ones that start at the boundary of the stopping region. The upper part of the boundary is absorbing, the left one is reflecting. For trajectories that pass via kink of the boundary, it is optimal to either stop when the beliefs reach the kink (time $T^* < \infty$) or wait until the next observation (time $T^* = \infty$).

In Fig. 5, for the same parameters of the model, we show the graphs of the value function and $T = T^*$. In a large part of the inaction region, $T^* = \infty$; this corresponds to $T = 0.5$ on the graph. In Fig. 6, we show the graphs of the same functions in a neighborhood of the absorbing part of the boundary.

5. Conclusion

We characterized equilibrium strategies in a two-player stopping time game with information externalities. The players experiment with profit generating technologies

which may experience costly breakdowns only the technology is of "bad" quality. In the latter case, breakdowns arrive at jump times of two standard Poisson processes with the same intensity. Initially, the type of technology is unknown, but it is the same type for both players. The costs of breakdowns are stochastic, and they evolve as a process independent of the Poisson processes. The Poisson processes may be viewed as idiosyncratic risk, and the cost process may be interpreted as an aggregate risk. Players impose positive information externality on each other since each player learns the current value of the cost when she incurs the cost herself, or when the other player incurs the cost. As a result, players may experiment longer than a single player would do. All uncertainty and payoff relevant information is public.

At each point in time, each player can make an irreversible stopping decision. We characterized the stopping region in the space of beliefs and formulated general conditions on the evolution of beliefs near the boundary of the exit region, which ensure that the value function is smooth at a part of the boundary (absorbing part), and the condition that ensure that the value function has a kink at a part of boundary (reflecting part).

The immediate extensions of this project are to consider the case of inconclusive breakdowns, when the "good" technology breaks down less frequently than the bad one as in Keller and Rady [28], or the case when the qualities of technologies are correlated as in Klein and Rady [29] or Rosenberg et al. [40]. Other extensions include private types as in Murto and Välimäki [37] or private payoffs as in Heidhues et al. [21] or Rosenberg et al. [40].

References

[1] L. Alili and A. Kyprianou, " A Some remarks on first passage of Lévy process, the American put and pasting principle", *Annals of Applied Probability*, 15 (2005), 2062-2080.

[2] G. Barles, J. Bardeaux, M. Romano, and N. Samsoen, "Critical stock price near expiration," *Mathematical Finance*, 5 (1995), 77-95.

[3] D. Bergemann, and U. Hege, "Dynamic venture capital financing, learning and moral hazard," *Journal of Banking and Finance*, 22 (1998), 703-735.

[4] D. Bergemann, and U. Hege, "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36:4 (2005), 719-752.

[5] D. Bergemann, and J. Välimäki, "Learning and strategic pricing," *Econometrica*, 64 (1996), 1125-1149.

[6] D. Bergemann, and J. Välimäki, "Experimentation in markets," *Review of Economics Studies*, 67 (2000), 213-234.

[7] D. Bergemann, and J. Välimäki, "Dynamic price competition," *Journal of Economic Theory*, 127 (2006), 232-263.

[8] D. Bergemann, and J. Välimäki, "Bandit problems." In The New Palgrave Dictionary of Economics (Steven N. Durlauf and Larry E. Blume, eds.), PalgraveMacmillan, Basingstoke (2008).

[9] P. Bolton, and C. Harris, "Strategic experimentation", *Econometrica* 67 (1999), 349-374.

[10] S.I. Boyarchenko and S.Z. Levendorskiĭ, "Models of investment under uncertainty when shocks are non-Gaussian. The impact of the policy uncertainty on investment", *Working Paper Series of EERC (Economics Education and Research Consortium - Russia)* No. 98-02E, 1998. Available at http://eercnetwork.com/paper?page=35

[11] S.I. Boyarchenko and S.Z. Levendorskiĭ, "Perpetual American options under Lévy processes", *SIAM Journal on Control and Optimization*, 40:6 (2002), 1663-1696.

[12] S.I. Boyarchenko and S.Z. Levendorskiĭ, "Non-Gaussian Merton-Black-Scholes Theory", *Adv. Ser. Stat. Sci. Appl. Probab.*, 9, World Scientific Publishing Co., River Edge, NJ, 2002.

[13] S.I. Boyarchenko and S.Z. Levendorksiĭ, "Practical guide to real options in discrete time", *International Economic Review* 48:1 (2007), 311-342

[14] S.I. Boyarchenko and S.Z. Levendorksiĭ, "Irreversible Decisions Under Uncertainty (Optimal Stopping Made Easy)", Springer, Berlin, 2007.

[15] S.I. Boyarchenko and S.Z. Levendorksiĭ, "Poisson bandits of evolving shade of gray," Working paper, available at SSRN: http://ssrn.com/abstract=2649713

[16] A. Cohen, and E. Solan, "Bandit Problems with Lévy Payoff Processes," Mathematics of Operations Research, 38 (2013), 92?107.

[17] D. Cummins, and P.M. Barrieu, "Innovations in insurance markets: hybrid and securitized risk." In Handbook of Insurance (Georges Dionne, ed.) Springer, 2nd ed. (2013)

[18] J.-P. Decamps, and T. Mariotti, "Investment timing and learning externalities", *Journal of Economic Theory* 118 (2004), 80-102.

[19] J.-P. Decamps, T. Mariotti, and S.Villeneuve, "Investment timing under incomplete information", *Mathematics of Operations Research* 30 (2005), 472-500.

[20] J.-P. Decamps, T. Mariotti, and S.Villeneuve, "Investment timing under incomplete information: Erratum", *Mathematics of Operations Research* 34 (2009), 255-256.

[21] P. Heidhues, S. Rady, and P. Strack, "Strategic experimentation with private payoffs," *Journal of Economic Theory*, 159 (2015), 531-551.

[22] J. Hörner, and A. Skrzypacz, "Learning, experimentation and information design," Working Paper, 2016.

[23] T.D. Gerarden, R.G. Newell, R.N. Stavins, and R.C. Stowe, "An assessment of the energy-efficiency gap and its implications for climate change policy," NBER Working Paper 20905 http://www.nber.org/papers/w20905

[24] J. Hörner, N.A. Klein, and S. Rady, "Strongly symmetric equilibria in bandit games," *Cowles Foundation Discussion Paper No. 1056* (2014), Available at SSRN: http://ssrn.com/abstract=2482335

[25] G. Keller, S. Rady, and M. Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73:1 (2005), 39-68.

[26] G. Keller, and S. Rady, "Optimal experimentation in a changing environment," *Review of Economic Studies*, 66 (1999), 475-507.

[27] G. Keller, and S. Rady, "Strategic Experimentation with Poisson Bandits," *Theoretical Economics*, 5 (2010), 275-311.

[28] G. Keller, and S. Rady, "Breakdowns," *Theoretical Economics*, 10 (2015), 175-202.

[29] N. Klein, and S. Rady, "Negatively correlated bandits," *Review of Economic Studies*, 78:2, 693-732.

[30] H. Kunreuther, and E. Michel-Kerjan, "Economics of natural catastrophe risk insurance." In the Handbook of the Economics and Uncertainty (Mark J. Machina, and W. Kip Viscusi, eds.), Elsevier (2014).

[31] D. Lamberton, "Critical price for an American option near maturity," in: *Seminar on Stochastic Analysis, Random Fields and Applications*, M. Dozzi and F. Russo, eds. Birkhäuser, Boston Basel Berlin (1995), 353-358.

[32] S. Levendorskiĭ, "Pricing of the American put under Lévy processes", Research Report MaPhySto, Aarhus, 44 (2002), Available at http://www.maphysto.dk/cgi-bin/gp.cgi?publ=441.

[33] S. Levendorskiĭ, "Pricing of the American put under Lévy processes, *International Journal of Theoretical and Applied Finance*, 7:3 (2004), 303–335.

[34] S. Levendorskiĭ, "American and European Options in Multi-Factor Jump-Diffusion Models, Near Expiry", *Finance and Stochastics* 12:4 (2008), 541-560.
[35] M. Ludkovski, and R. Sircar, "Exploration and Exhaustibility in Dynamic Cournot Games," *European Journal of Applied Mathematics*, 23 (2012), 343-372.
[36] R.C. Merton, "Option Pricing when Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3 (1976), 125-144.
[37] P. Murto, and J. Välimäki, "Delay and information aggregation in stopping games with private information," *Journal of Economic Theory*, 148 (2013), 2404-2435.
[38] G. Peskir, and A. Shiryaev, *Optimal Stopping and Free-Boundary problems.* Basel: Birkhäuser Verlag.
[39] K. Roberts, and M.L. Weitzman, "Funding criteria for research, development, and exploration projects," *Econometrica*, 49:5 (1981), 1261-1288.
[40] D. Rosenberg, A. Salomon, and N. Vieille, "On games of strategic experimentation," *Games and Economic Behavior*, 82 (2013), 31-51.
[41] M. Rothschild, "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 9:2 (1974), 185-202.
[42] M.L. Weitzman, "Optimal search for the best alternative," *Econometrica*, 47:3 (1979), 641-654.

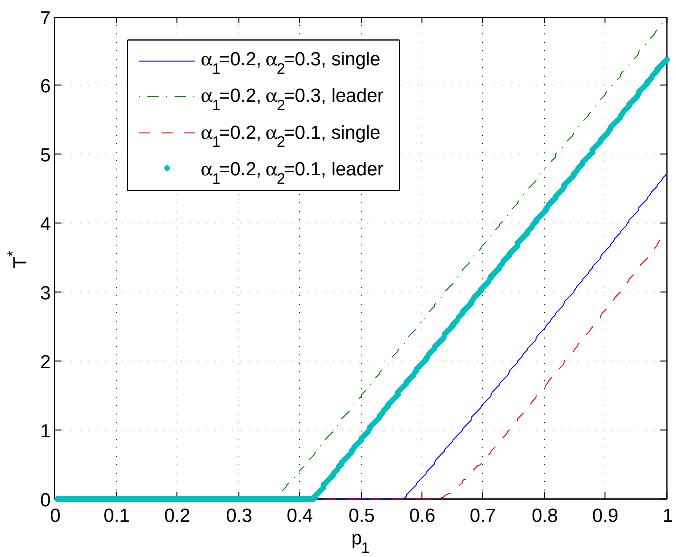


Figure 1. Value function of a single player and value of the same player when the other player precommits not to exit earlier, for two sets (α_1, α_2). Parameters: $\lambda = 0.8$, $r = 0.1$, $\mu_1 = 0.8$, $\mu_1 = 1.2$, $R = 30.3$, $c_l = 20$, $c_m = 50$, $c_h = 65$.

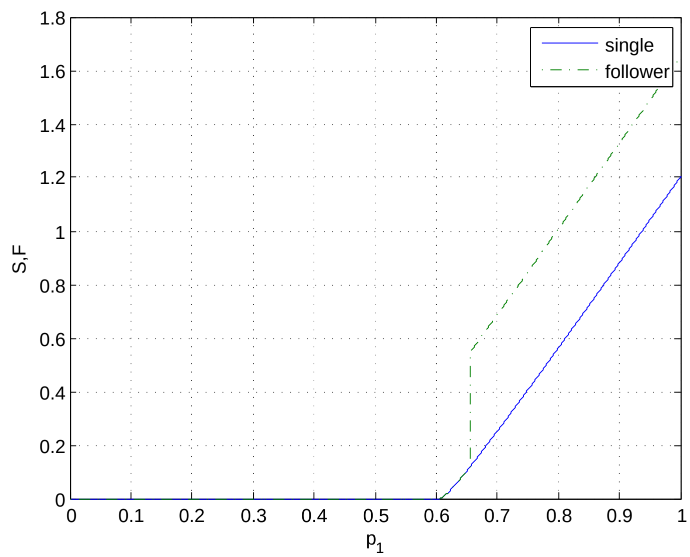
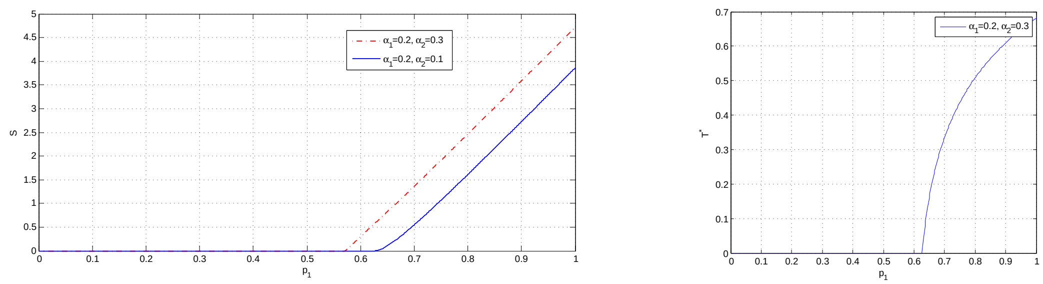


FIGURE 2. Value function of the follower. Parameters: $\lambda = 1$, $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $r = 0.1$, $\mu_1 = 4$, $\mu_1 = 6$, $R = 39$, $c_l = 20$, $c_m = 50$, $c_h = 65$.



(A) (B)

FIGURE 3. Conclusive breakdowns, value function of a single player (A), and exercise strategy (B). Parameters: $R = 30$, $\lambda = 0.8$, $r = 0.1$, $\mu_1 = 0.8$, $\mu_1 = 1.2$, $c_l = 20$, $c_m = 50$, $c_h = 65$.

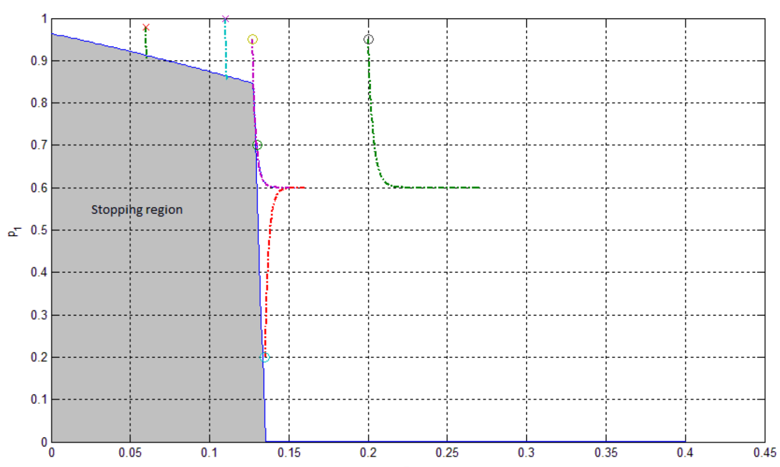
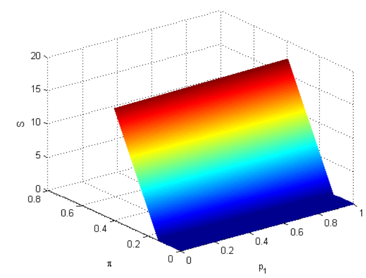
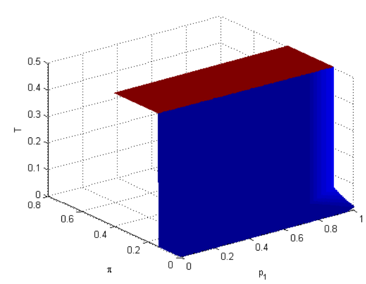


FIGURE 4. Stopping region and typical trajectories. Crosses: starting beliefs for trajectories that enter the stopping region ($T^* < \infty$). Circles: starting beliefs for trajectories that never enter the stopping region ($T^* = \infty$) Parameters: $\lambda = 0.2$, $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $r = 0.1$, $\mu_1 = 4$, $\mu_1 = 6$, $R = 5.5$, $c_l = 20$, $c_m = 50$, $c_h = 65$.

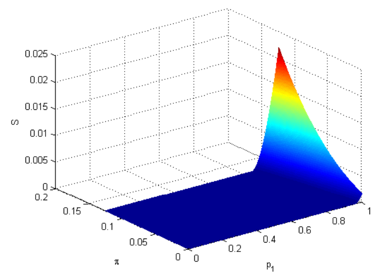


(A) Graph of value function.

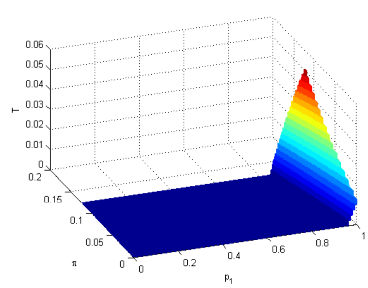


(B) Graph of $T = T^*$.

FIGURE 5. Parameters: $\lambda = 0.2$, $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $r = 0.1$, $\mu_1 = 4$, $\mu_1 = 6$, $R = 5.5$, $c_l = 20$, $c_m = 50$, $c_h = 65$.



(A) Graph of value function.



(B) Graph of $T = T^*$.

Figure 6. Parameters: $\lambda = 0.2$, $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $r = 0.1$, $\mu_1 = 4$, $\mu_1 = 6$, $R = 5.5$, $c_l = 20$, $c_m = 50$, $c_h = 65$.