STRATEGIC EXPERIMENTATION WITH HUMPED BANDITS

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

Abstract. Risks related to events that arrive randomly play important role in many real life decisions, and models of learning and experimentation based on two-armed Poisson bandits addressed several important aspects related to strategic and motivational learning in cases when events arrive at jump times of the standard Poisson process. At the same time, these models fail to explain some interesting features of reality. We suggest a new class of models of strategic experimentation which are almost as tractable as exponential models, but incorporate such realistic features as dependence of the expected rate of news arrival on the time elapsed since the start of an experiment and judgement about the quality of a risky arm based on evidence of a series of trials as opposed to a single evidence of success or failure as in exponential models with conclusive experiments. We show that, unlike in the exponential models, players may stop experimentation before the first failure happens. We also demonstrate a crowding out effect in models with profitable breakthroughs.

Keywords: stopping time games, time-inhomogeneous Poisson process, strategic experimentation, breakdowns, breakthroughs

JEL: C73, C61, D81

1. Introduction

1.1. **Motivation.** This paper suggests a new class of learning and experimentation models based on *humped bandits*. The proposed models are almost as tractable as models based on Poisson bandits and generate qualitatively new results. The models based on humped bandits can used as a framework for experimentation related to

I am thankful for discussions to participants of research seminars at Toulouse School of Economics, Carlos III University of Madrid, Tilburg University, The Center for Rationality in the Hebrew University of Jerusalem, at the University of Tel Aviv, at Haifa University, at Texas A&M University, at the Research University Higher School of Economics, Moscow and St. Petersburg. I am especially grateful to Ehud Lehrer and Eilon Solan for pointing out mistakes in the first version of the paper and making useful suggestions and to William Fuchs for suggestion to consider a more general class of strategies. I am also thankful for discussions to Adrien Blanchet, Jean-Paul Decamps, William Fuchs, Ángel Hernando-Veciano, Kuno Huisman, Ilan Kremer, Peter Kort, Rida Laraki, Amnon Maltz, Abraham Neyman, Frank Riedel, Anna Rubinchik, Jan-Henrik Steg, Max Stinchcombe, Nicolas Vieille, and Eyal Winter. The usual disclaimer applies.

safety of medications or pollutants that have a potential for accumulation and long term storage in a human body; for timing exit decisions of financial institutions dealing with risky assets whose riskiness may accumulate over time; for designing grant competitions for long term projects; or for determining optimal terms of politicians.

In many situations in real life, it is necessary to quantify risks related to events that arrive at random times, as well as frequency of their arrivals. For example, in finance, it is necessary to evaluate default risks of borrowers or assets and default rates. In pharmaceuticals, it is necessary to evaluate possible side effects of a new drug or its efficiency. In any kind of sponsored research, sponsors have to figure out the probability of success, and so on. However, small chances of arrivals of random events may increase as time goes by. For example, strong hurricanes are expected to happen more and more frequently as a result of the global warming. Chemicals that have an ability to deposit in areas of human body that are rich in fat are more likely to impair health after they have been used for a longer time. A risky borrower is less likely to default immediately upon the loan initiation than some time after. When a researcher starts a new research project, sponsors hardly expect her to generate immediate success even when they have high beliefs in the quality of the project. All aforementioned situations can be modeled if one assumes that the rate of arrival of random events increases in time.

Suppose now, that *a priory*, it is not known whether the rate of arrival of future events in some of the above examples is zero or positive. For example, a new financial instrument may never default, or it may default at a random time in the future. In this case, will it be optimal for a financial institution to keep using this instrument until a default happens (if ever) or stop using it before anything happens? If one assumes that the rate of arrival of default remains constant independently of how long the asset has been used, then it is optimal to use it until the default. However, if one assumes that the rate of arrival of default is an increasing function of time, it may become optimal to get rid of the asset before anything happens. For example, in February 2007, just before everything fell apart, Goldman Sachs sold thousands of subprime mortgages to investors. In July 2016, the Fixed Income Clearing Corporation (FICC) filed with the Securities and Exchange Commission (SEC) a proposal to suspend interbank service of the General Collateral Finance (GCF) Repo service. One of the major post-crisis innovations to the Collateralized Debt Swaps (CDS) market was the introduction of central clearing for certain types of contract. In the US, index (CDX) contracts are mandatory cleared on ICE Clear and the Chicago Mercantile Exchange (CME). ICE started clearing index contracts in March 2009, and CME started clearing index contracts in December 2009. Clearing involves risk, so in September 2017 CME announced exit from CDS clearing business.

1.2. **Related literature.** There is an extensive literature on learning and experimentation based on so called two or muti-armed bandits. A standard "two-armed" bandit is an attempt to describe a hypothetical experiment in which a player faces two slot machines; the quality of one of the slots is known (safe arm), and the other one (risky arm) may be good or bad. In case of so called "conclusive" experiments - the first event observed on the risky arm reveals its quality completely, so the experiment is over, when the first ("conclusive") success or failure is observed. 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., [1, 2, 3, 4, 5, 6, 20, 28, 30, 31] and references therein).

Models of strategic learning and experimentation extend "two-armed" bandit experiments to a setting where several players face copies of the same slot machine. Players then learn about the quality of the risky arm not only from outcomes of their own experiments, but also from their colleagues. In models of strategic experimentation, it is common to assume away payoff externalities and focus on information externalities, the role of information, and, in more advanced settings, on design of information. See, for example, [7, 10, 18, 19, 21, 22]. Recent developments include (but are not limited to) correlated risky arms as in Klein and Rady [23] and Rosenberg et al. [29], or private payoffs as in Heidhues et al. [15] and Rosenberg et al. [29], or departures from Markovian strategies as in Hörner et al. [18]. For other developments and an excellent comprehensive review of the literature see Hörner and Skrzypacz [16] and references therein.

In continuous time models, the payoff generated by the risky arm of a "two-armed" bandit follows a certain continuous time stochastic process whose parameters are not known. For example, Bolton and Harris [7] model the unknown payoff as a Brownian motion with unknown drift and known variance in a model of strategic experimentation. Decamps et al. [11, 12] 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. [19], Keller and Rady [21, 22] use a Poisson process with unknown rate of arrival to model the risky arm. Decamps and Mariotti [10] study a duopoly model of investment where a signal about the quality of the project is modeled as a Poisson process. Cohen and Solan [8] 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.

There are typically, two kinds of Poisson bandits models: bad news models and good news models. Bad news are costly failures, and good news are profitable successes; the type of news is know before an experiment starts. The main feature of exponential bandits models is that the rate of arrival of news over a time interval $(t, t + \Delta t)$ is independent of t if the quality of the risky arm is known. This is convenient for tractability, but not very realistic.

As a result of the above assumption, the rate of arrival of news is decreasing in time, if the quality of the risky arm is unknown. If nothing happens in the model with potentially bad news, players become more and more optimistic about the quality of the risky arm. Therefore, if it is optimal to start experimentation at the prior beliefs levels, experimentation never stops until the first breakdown occurs (if ever). On the contrary, if no successes arrive in the model with potentially good news, players become more and more pessimistic about the quality of the risky arm, and experimentation always stops in finite time, unless the first success arrives earlier. Due to the above, players experiment too long in bad news models, and experimentation levels in good news models are too low.

The papers which are mostly close to our paper are Keller et al. [19] Keller and Rady [21, 22] and Rosenberg et al. [29]. Keller et al.[19] and Keller and Rady [21] study good news model, and [22] study bad news model; in either model, news arrive at the jump times of Poisson processes which are independent. Rosenberg et al. [29] consider an irreversible exit problem in a model with breakthroughs with correlated risky arms both in the case when payoffs are public and private.

1.3. **Overview of the results.** We propose a class of experimentation models where the rate of arrival of news is increasing in time, if the quality of the risky arm is known. If the quality of the risky arm is unknown, the rate of arrival of news becomes hump-shaped. Right after the experiment starts, the rate of arrival increases in time until it reaches a certain maximal level; then it starts decreasing and behaving more and more as a negative exponential as the experiment "grows older." The longer the experimenter observes no bad news, the more optimistic she becomes about the quality of the risky arm, but at the same time, if the risky arm is bad, the rate of arrival of a failure grows, and these two opposite movements may make stopping before the first breakdown optimal. If such stopping is optimal, it happens where the rate of arrival of news is increasing. As opposed to this, in a model with breakthroughs, experimentation does not stop while the rate of arrival of good news keeps growing, even though experimenters become more pessimistic about the quality of the risky arm. Khan and Stinchcombe [24] find similar results in semi-Markovian decision theory. Namely, they identify two classes of situations in which delay in decision systems is optimal: in the first class delay is optimal when the hazard rate of further changes is increasing, and in the second class, delay is optimal when the hazard rate is decreasing.

We show that, in the absence of news, it is not optimal to stop experimentation until the rate of arrival of news reaches a certain critical level. In a model with failures, a player stops experimentation when the marginal benefit from staying active equals the expected marginal cost. Stopping prior to the first arrival of bad news is a new result, which makes our model qualitatively different from the Poisson bandits models.

If the marginal benefit of staying active is higher than the expected marginal cost at any time, the player never stops before the first failure is observed (if ever). In a model with breakthroughs, a player stops experimentation when the expected marginal benefit from staying active equals the marginal cost. If parameters of a good and bad news models are such that the cutoff expected rates of the news arrival are the same, experimentation in the good news model lasts longer than in the bad news model, provided no news arrive in either model, which is another qualitative difference from experimentation models based on Poisson bandits.

We show that in generic cases in the bad news model, in equilibrium, all players either do not stop until the first failure is observed, or stop simultaneously before the first failure. While the critical level that the hazard rates have to rich before the players stop is independent of the number of players, it takes longer (respectively, shorter) to rich this level as the number of players increases in the bad news model (respectively, in the good news model). In either case, the critical values of the hazard rates are the same as in case of cooperative experimentation - from this perspective, there is no loss of efficiency in strategic experimentation. Another interesting effect which can be observed in our model, and it is absent in the corresponding exponential model is *crowding out*. For each critical value of the hazard rate in the good news model, there exists a critical number of players, say, n^*, s.t. experimentation is profitable for $n < n^*$ players and not profitable for $n \geq n^*$ players. Existence of the crowding out effect may have interesting policy implications for designers of grant competitions, innovation contests etc.

The rest of the paper is organized as follows. The primitives of the experimentation model are decribed in Section 2. The stopping game is described in Section 3. Section 4 considers a stopping time game in a bad news model. The detailed analysis of the game and construction of subgame perfect equilibria are in Section 5. A stopping game in a good news model is analyzed in Section 6. Generalizations for more general intensities of the news arrival are outlined in Section 7. Section 8 concludes. Technical proofs are relegated to the appendix.

2. Model decription

2.1. **The setup.** Time $t \in \mathbb{R}_+$ is continuous, and the discount rate is $r > 0$. Let n symmetric players experiment with risky projects of unknown quality, such as a nuclear technology, a defaultable loan, or a new drug. The quality of a project is identified with its ability to generate *news* (or observations). The quality depends on the state of the nature denoted by $\theta \in \{0, 1\}$. If $\theta = 0$, the project generates no news; if $\theta = 1$, the project generates news that arrive at random times of a time-inhomogeneous Poisson process. News can be good (profitable breakthroughs) or bad (costly breakdowns). In the real life the nature of news may not be known in advance, for instance, a drug developer may not know if the drug will be effective to cure a targeted decease of cause serious side effects. In this paper, we follow a tradition in the bulk of experimentation literature and assume that the type of news is known before the experimentation starts.

We leave for the future study the case when the players can have projects of different types, and the types may be positively of negatively correlated as in Rosenberg et al. [29] or Klein and Rady [23], and assume that the quality of the projects is the same for all players, so that if one of the players receives a piece of news, the other players conclude that their projects generate the same type of news for sure. Processes for news arrival are independent of each other, but have the same characteristics. The initial common prior assigns probability $\bar{\pi} \in [0,1]$ to $\theta = 1$.

2.2. **Uncertainty.** In this Section, we specify the model primitive, evolution of beliefs and rates of arrival of news. We also characterize properties of beliefs and hazard rates that will be used later in construction of equilibria. The primitive of the model is the rate of arrival (hazard rate) of news at time t of a single project when $\bar{\pi} = 1$. Denote this hazard rate by $\Lambda_1(1;t)$. Here and below the subscript indicates the number of active projects/players, and the first argument specifies the prior belief. Assume that the hazard rate satisfies the following properties:

(i) $\Lambda_1(1;0) = 0$;
(ii) $\Lambda_1(1;t)$ is continuous and increasing in t.

The first property simplifies some of the technical proofs, but it will be relaxed in Section 7, where alternative models for the hazard rate will be presented. The second property introduces a natural dependence of the rate of arrival on the time which has elapsed since the start of an experiment. As the leading example, we use

$$\Lambda_1(1;t) = \frac{\lambda^2 t}{1+\lambda t}, \tag{2.1}$$

where $\lambda > 0$. This hazard rate corresponds to the case when the time of news arrival is $Erlang(2,\lambda)$ random variable[1]. Note that, in this case, the expected time until the first observation is $2/\lambda$.

If $\bar{\pi} = 1$, the probability of the event that no news is observed before t by a single experimenter is

$$p_1(1;t) = \exp\left(-\int_0^t \Lambda_1(1;s)ds\right). \tag{2.2}$$

Clearly, $p_1(1;0) = 1$, and by property (ii), $\lim_{t\to\infty} \int_0^t \Lambda_1(1;s)ds = +\infty$, therefore $\lim_{t\to\infty} p(1;t) = 0$. For the leading example, $p_1(1;t) = (\lambda t + 1)e^{-\lambda t}$.

If $\bar{\pi} < 1$, the probability of the event that no news is observed before t by one experimenter is

$$p_1(\bar{\pi};t) = 1 - \bar{\pi} + \bar{\pi} p_1(1;t). \tag{2.3}$$

Obviously, $p_1(\bar{\pi};0) = 1$, and $\lim_{t\to\infty} p_1(\bar{\pi};t) = 1-\bar{\pi}$. Furthermore, $p_1(\bar{\pi};t)$ is decreasing in t, because

$$\partial_t p_1(\bar{\pi};t) = \bar{\pi}\partial_t p_1(1;t) = -\bar{\pi}\Lambda_1(1;t)p_1(1;t) < 0 \;\forall\, t > 0; \tag{2.4}$$

here and below ∂_t denotes the partial derivative w.r.t. the second argument.

If n players experiment, the probability of no news arriving before time t is

$$p_n(\bar{\pi};t) = 1 - \bar{\pi} + \bar{\pi} p_1(1;t)^n. \tag{2.5}$$

[1] The p.d.f. of $Erlang(k,\lambda)$ distribution for $k \geq 1$ and $\lambda > 0$ is given by $f(t) = \lambda^k t^{k-1} e^{-\lambda t}/(k-1)!$.

If no news were observed until time t, the posterior belief assigns the probability $\pi_n(\bar{\pi};t)$ to the event $\theta = 1$, and, by the Bayes rule,

$$\pi_n(\bar{\pi};t) = \frac{\bar{\pi} p_1(1;t)^n}{p_n(\bar{\pi};t)}. \tag{2.6}$$

Straightforward differentiation shows that

$$\partial_t \pi_n(\bar{\pi};t) = -n\Lambda_1(1;t)\pi_n(\bar{\pi};t)(1-\pi_n(\bar{\pi};t)) < 0 \; \forall\, t > 0. \tag{2.7}$$

The posterior belief is decreasing in t - the more time passed without any observations, the lower is the probability that the risky project will generate news.

Lemma 2.1. *Let $n > 1$. For any $t' > 0$,*

(i) $\pi_n(\bar{\pi};t') < \pi_{n-1}(\bar{\pi};t')$;

(ii)

$$\pi_n(\bar{\pi};t') = \pi_{n-1}(\pi_1(\bar{\pi};t'), t'); \tag{2.8}$$

(iii) functions $\pi_n(\bar{\pi};t)$ and $\pi_{n-1}(\pi_1(\bar{\pi};t'), t$ intersect only at $t = t'$; and

(iv)

$$\pi_n(\bar{\pi};t') < \pi_{n-1}(\pi_1(\bar{\pi};t'), t)) \;\Leftrightarrow\; t > t'. \tag{2.9}$$

The first statement says that any level of beliefs will be reached faster if more players experiment, starting with the same prior. The second and third statements say that if one of the n players stops experimentation at t', the rest $n-1$ players continue with the starting belief $\pi_1(\bar{\pi};t')$, and this new belief is unique. The last statement says that beliefs of the remaining $n-1$ players decrease slower than beliefs of n players if none of the n players had stopped. See Section 9.1 for the proof.

Next, we define the rate of arrival of observations at time t when $n \geq 1$ players experiment with the projects of unknown quality:

$$\Lambda_n(\bar{\pi};t) = n\Lambda_1(1;t)\pi_n(\bar{\pi};t).$$

Since the players are symmetric, each of them is equally likely to receive the first news, therefore the individual hazard rate is

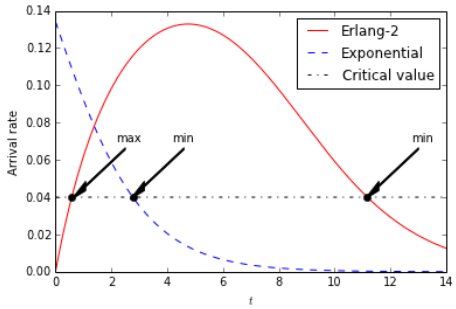
$$\frac{1}{n}\Lambda_n(\bar{\pi};t) = \Lambda_1(1;t)\pi_n(\bar{\pi};t). \tag{2.10}$$

Since $\pi_n(\bar{\pi};t) < \pi_{n-1}(\bar{\pi};t)$ for all $t > 0$,

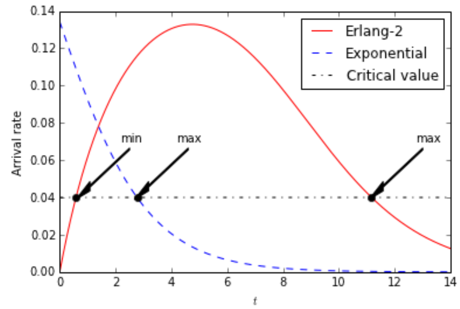
$$\frac{1}{n}\Lambda_n(\bar{\pi};t) < \frac{1}{n-1}\Lambda_{n-1}(\bar{\pi};t) \; \forall\, t > 0. \tag{2.11}$$

The next result describes changes in the hazard rates when the first of n players stops experimentation.

Lemma 2.2. *Let $\Lambda_1(1;t)$ satisfying the standing assumptions (i)-(ii) be of the class C^1 w.r.t. the second argument, let there exist $\partial_t\Lambda_1(1;+0)$, and let it be finite. Then*



(A) Bad news model



(B) Good news model

FIGURE 1. Erlang-2 vs. exponential model

(a) for any $t' > 0$, *functions* $\frac{1}{n}\Lambda_n(\bar{\pi}; t)$ *and* $\frac{1}{n-1}\Lambda_{n-1}(\pi_1(\bar{\pi}; t'); t)$ *intersect only at* $t = 0$ *and* $t = t'$*; and*

(b)

$$\frac{1}{n}\Lambda_n(\bar{\pi}; t) < \frac{1}{n-1}\Lambda_{n-1}(\pi_1(\bar{\pi}; t'), t) \;\Leftrightarrow\; t > t'. \tag{2.12}$$

See Section 9.2 for the proof. The following result describes the shape of the hazards rates.

Lemma 2.3. *Let* $\Lambda_1(1; t)$ *satisfying the standing assumptions (i)-(ii) be of the class* C^2 *w.r.t. the second argument, and let also the following properties hold:*

(iii) there exists $\partial_t \Lambda_1(1; +0)$, *and it is finite;*

(iv)

$$\partial_t^2 \left(\frac{1}{\Lambda_1(1; t)} \right) > 0, \;\; \forall\, t > 0;$$

(v) and

$$\lim_{t \to \infty} \partial_t \left(\frac{-1}{\Lambda_1(1; t)} \right) < 1.$$

Then as a function of t, $\Lambda_n(\bar{\pi}; t)$ *has the global maximum* $\hat{t}_n = \hat{t}_n(\bar{\pi})$.

Lemma 2.3 states that the hazard rate $\Lambda_n(\bar{\pi}; t)$ is hump-shaped. In the exponential model, the corresponding hazard rate $\Lambda_n^{\exp}(\bar{\pi}; t) = n\lambda\pi_n(\bar{\pi}; t)$ is proportional to the beliefs. Since the beliefs are decreasing in time, the hazard rate is also decreasing.

Exponential bandits models use the beliefs at time t as the state variable and formulate optimal stopping strategies in terms of critical beliefs. In our model, stopping strategies have to be formulated in terms of critical hazard rates, and the hazard rate at time t is a natural state variable. Let $A > 0$ denote the critical value. If $A \geq \Lambda_n(\bar{\pi}; \hat{t}(\bar{\pi}))$, the players either never start experimentation (in case of the good news model) or never stop before the first failure arrives (in the bad news model). If

$A < \Lambda_n(\bar{\pi}; \hat{t}(\bar{\pi}))$, then the equation

$$A = \Lambda_n(\bar{\pi}; t)$$

has two solutions. In the bad news model, the smaller solution is the local maximum and the larger solution is the local minimum of the value function of a player. In the good news model, the smaller solution is the local minimum and the larger solution is the local maximum of the value function of the player. As opposed to this, equation

$$A = \Lambda_n^{\exp}(\bar{\pi}; t)$$

has at most one solution, which is the local minimum of the value function in the bad news model, and the local maximum of the value function in the good news model. See Fig. 1 for illustration of the leading example. That is why in the exponential model, players never stop before the first breakdown happens. In our model, it may be optimal to stop before the first failure. Later, we provide conditions on the model primitives which ensure that the local maximum of the value function is also the global maximum.

Proof. Rewrite the derivative (9.2) as

$$\frac{1}{n}\partial_t \Lambda_n(\bar{\pi}; t) = \Lambda_1(1; t)^2 \pi_n(\bar{\pi}; t) \left[\frac{\partial_t \Lambda_1(1; t)}{\Lambda_1(1; t)^2} - n(1 - \pi_n(\bar{\pi}; t))\right].$$

Observe that the function

$$f(t) := \frac{\partial_t \Lambda_1(1; t)}{\Lambda_1(1; t)^2} = -\partial_t \left(\frac{1}{\Lambda_1(1; t)}\right)$$

is strictly decreasing due to assumption (iv); $\lim_{t \to +0} f(t) = +\infty$ by assumptions (i) and (iv); and $\lim_{t \to \infty} f(t) < 1$ by assumption (v). The function $\phi(t) := 1 - \pi_n(\bar{\pi}; t)$ is strictly increasing; $\phi(0) = 1 - \bar{\pi}$, and $\lim_{t \to \infty} \phi(t) = 1$.

The sign of the derivative $\partial_t \Lambda_n(\bar{\pi}; t)$ is determined by the sign of the difference $f(t) - \phi(t)$. The difference is a continuous decreasing function which is positive in a right neighborhood of zero and negative for sufficiently large t. Hence there exists a unique $\hat{t}_n = \hat{t}_n(\bar{\pi}) > 0$ s.t. $f(\hat{t}_n(\bar{\pi})) = \phi(\hat{t}_n(\bar{\pi})$, and

$$f(t) > \phi(t) \;\Leftrightarrow\; t < \hat{t}_n(\bar{\pi}).$$

Hence, $\hat{t}_n(\bar{\pi})$ is the global maximum of $\Lambda_n(\bar{\pi}; t)$. □

3. Stopping game

In this Section, we will focus on the case of two players. By induction, the results can be extended to the case $n > 2$ in a straightforward manner.

3.1. **Strategies.** We consider the game of timing, characterized by the following structure. Two players experiment with projects of unknown quality as described in Section 2. All payoffs, function $\Lambda_1(1;t)$, and the players' actions are public information. W.l.o.g. assume that the game starts at $t = 0$. At each point $t \geq 0$, player $i \in \{1, 2\}$ may make an irreversible stopping decision conditioned on the history of the game.

At any $t \geq 0$, the history of the game includes observations of news arrivals (including the empty set if no observations arrived 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 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 player i got news for the s^{th} time. The history of observations at any $t \geq 0$ is

$$O_t = \{(\tau_1^{s'})\}_{s' \leq t \wedge T_1} \cup \{(\tau_2^{s''})\}_{s'' \leq t \wedge T_2}.$$

A typical history h at time t is $h_t = (O_t, \tilde{t}_1(t), \tilde{t}_2(t))$. If $T_i < T_j$, we call player i the leader, and player j the follower.

For simplicity, we will consider the case, when experimentation stops after the first observation. This supposition can be justified either by assumption that learning the true quality of the project is the only objective of the players, or by an appropriate specification of the payoff functions. In particular, in the bad news model, we assume that an active player gets a flow payoff $rR > 0$ as long as no failure was observed. This stream can be viewed, for example, as sponsored research contributions, or revenue generated by a project net of insurance costs, or mortgage payments. If the project is bad, then after the first failure, the stream of revenues disappears (e.g., the sponsor withdraws support from a pharmaceutical company as soon as a side effect of a new drug is observed; the insurance company increases the premium to the extent that offsets the revenue stream of a faulty technology; a borrower is not able to make monthly payments after a default, etc). Given this assumption, experimentation after the first failure becomes non-profitable, so the players stop experimenting as soon as they learn that the quality of the project is bad. In the good news model, we assume that as soon as the first success is observed, the player who achieved the success wins the prize, the other experiment participants get the value of the outside option, and the game is over.

Due to the above assumption, if $O_t \neq \emptyset$, the game is over. Since any agent which remains in the game exits at the moment of the first observation, we may define the strategies only for the histories of the form $h_t(\emptyset, \tilde{t}_1(t), \tilde{t}_2(t))$. Denote by $(\Omega^j, \mathcal{F}^j, \{\mathcal{F}^j_s\}_{s=0}^\infty)$ the filtered measure space generated by T_j and $\tau = \tau_1^1 \wedge \tau_2^1$.

Definition 3.1. A strategy for player $i \in \{1, 2\}$ in the game starting at $t = 0$ is a process q_i taking values in $[0, 1]$, adapted to the filtration $\{\mathcal{F}^j_s\}_{s=0}^\infty$, $j \neq i$, with non-increasing, left-continuous with right limits (LCRL) trajectories; and $q_i(0) = 1$.

Denote by $\mathcal{H}_t$ the set of histories up to time t, s.t. $O_t = \emptyset$. Following Laraki et al. [25], and Dutta and Rustichini [13], for any time $t > 0$, define a proper subgame as the timing game that starts at the end of the history $h \in \mathcal{H}_t$.

Definition 3.2. A strategy for player $i \in \{1, 2\}$ in a subgame starting at $t > 0$, after a history $h \in \mathcal{H}_t$, is a process $q_i^{t,h}$ taking values in $[0, 1]$, adapted to the filtration $\{\mathcal{F}^j_s\}_{s=t}^\infty$, $j \neq i$, with non-increasing, left-continuous with right limits (LCRL) trajectories; and $q_i^{t,h}(t) = 1$.

Note that $q_i(t)$ is the probability that player i did not exit at time t or earlier, conditioned on no observations up to time t, and $q_i^{t,h}(s)$ is the probability that player i did not exit during $[t, s]$, conditioned on being active at t and no observations up to time s, given the history $h \in \mathcal{H}_t$. In either of the definitions above, we allow for the case $q_i(+\infty) > 0$ and $q_i^{t,h}(+\infty) > 0$, which means that player i decides not to exit ever with the positive probability $q_i(+\infty)$ and $q_i^{t,h}(+\infty)$, respectively, unless a failure happens. Thus, the probability that the player will not exit is $q_i(+\infty)\bar{\pi}$ and $q_i^{t,h}(+\infty)\bar{\pi}$, respectively.

Definition 3.3. A strategy of player i is called consistent, if for any $0 \le t \le t' \le s$, and $h \in \mathcal{H}_t$ and $h' \in \mathcal{H}_{t'}$ such that $h'|_{[0,t]} = h$,

$$q_i^{t,h}(s) = q_i^{t,h}(t') q_i^{t',h'}(s). \tag{3.1}$$

With slight abuse of notation, we write q_i^t instead of $q_i^{t,h}$.

3.2. **Value functions and equilibrium.** In order to define value functions of the players in this game, we will use the following version of the definition of the Riemann-Stieltjes integral.

Definition 3.4. Assume that the following conditions hold

(i) $q, \Psi : [0, +\infty) \to \mathbb{R}$ are bounded LCRL functions;
(ii) q is of finite variation;
(iii) the singular continuous component of the Lebesgue decomposition of q is trivial;
(iv) I is an interval of one of the following forms: $(a, b), [a, b), (a, b], [a, b], [a, +\infty), (a, +\infty)$, where $0 \le a < b < +\infty$ or a union of non-intersecting intervals of this form.

Define

$$\int_I \Psi(s) dq(s) = \int_I \Psi(s) q'(s) ds + \sum_{t_j \in I : \Delta q(t_j) \neq 0} \Psi(t_j) \Delta q(t_j), \tag{3.2}$$

where $\Delta q(t_j) := q(t_j + 0) - q(t_j)$.

Let $G_i(\bar{\pi};t)$ denote the instantaneous expected payoff flow of player i if none of the players stopped until time $t>0$. Let $F_i(\bar{\pi};t)$ denote the expected time t value of player i if player j stopped at time t, and player i did not. Finally, let $S\ge 0$ denote the value of the outside option. From now on, we will consider the simple strategies q_1,q_2, whose singular continuous components of the Lebesgue decompositions are trivial. Given the strategy profile (q_i,q_j), the value of player i in the game that starts at $t=0$ is

$$\begin{aligned} V_i(\bar{\pi};q_i,q_j) &= \int_0^\infty e^{-rt}p_2(\bar{\pi};t)G_i(\bar{\pi};t)q_i(t)q_j(t)dt \\ &\quad + \int_{\{t\ge 0 \ \mid\ \Delta q_i(t)=0\}} e^{-rt}p_2(\bar{\pi};t)F_i(\bar{\pi};t)q_i(t)(-dq_j(t)) \\ &\quad + \int_0^\infty e^{-rt}p_2(\bar{\pi};t)Sq_j(t)(-dq_i(t)). \end{aligned} \tag{3.3}$$

Here the first integral is the expected present value of the flow payoffs $G_i(\bar{\pi};t)$ conditioned on both players being active (which happens with probability $q_i(t)q_j(t)$) and on no news arrival before time t (which happens with probability $p_2(\bar{\pi},t)$). The second integral is the the expected present value of the payoff that player i gets at time t, provided this player is still active at t, player j stopped at t, and nothing else happened before t. Finally, the last integral is the expected present value of the outside option if player i exits at time t, provided no observations arrived before t. Later we will show that $G_i(\bar{\pi};\cdot)$ and $F_i(\bar{\pi};\cdot)$ are continuous and have finite limits as $t\to\infty$, hence, $V(\bar{\pi};q_i,q_j)$ is well-defined and finite. Note that the second integral in (6.2) takes into account jumps in q_j only, and the last integral takes into account jumps in q_i only as well as simultaneous jumps in q_i and q_j.

Definition 3.5. A strategy profile $\hat{q}=(\hat{q}_i,\hat{q}_j)$ is a Nash equilibrium for the game starting at $t=0$, if for every $(i,j)\in\{(1,2),(2,1)\}$

$$V_i(\bar{\pi};\hat{q}_i,\hat{q}_j)=\sup_{q_i} V_i(\bar{\pi};q_i,\hat{q}_j).$$

A profile of consistent strategies $\hat{q}^t=(\hat{q}_i^t,\hat{q}_j^t)$ is a subgame perfect Nash equilibrium (SPE) if for every $t\ge 0$, $\hat{q}^t$ is a Nash equilibrium in the subgame that starts at t (when payoffs are discounted to time t).

4. Bad news model

An active player gets a flow payoff $rR>0$ as long as no failure was observed, and zero after the first failure. In case of a failure, the player has to pay a lump-sum cost $C>0$ - for example, a new drug developer has to pay patients if they developed serious side effects while trying the new drug. Since the rate of arrival of news is zero at $t=0$, and remains small in a right neighborhood of zero, it is always optimal to start experimentation in the bad news model; and none of the players has yet stopped at the start of the game. As the rate of arrival $\Lambda_2(\bar{\pi};t)$ increases, it may become optimal for one or both players to quit. We will prove that, depending on parameters of the model, either the players do not stop until the first failure happens, or the stopping rules in pure strategies are of the threshold type - the players quit when the corresponding rates of arrival reach a certain threshold from below. The former outcome is possible only if the ratio $r(R-S)/C$ is sufficiently high.

Once one of the players has quitted experimentation, the other player faces a non-strategic stopping problem, which can be easily solved. Thus, when considering subgame perfect equilibria, we will first examine subgames when one of the players has stopped, and then move to subgames where neither player has quitted as yet. To simplify the notation, we suppress the dependence of value functions on the other player's strategy. Since the players are symmetric, we also drop the subscripts identifying the players in the follower's problem below..

4.1. **Follower's problem.** Consider a subgame that starts after the history such that no observations arrived, and only one of the players has stopped. Suppose, this happened at time t. Then the remaining player (the follower) chooses a strategy q_f^t satisfying conditions of Definition 3.2, which solves the following problem:

$$\begin{aligned} F(\bar{\pi};t) &= \sup_{q_f^t}\left[\int_t^\infty e^{-r(t'-t)}q_f^t(t')\frac{p_1(\pi_1(\bar{\pi};t);t')}{p_1(\pi_1(\bar{\pi};t);t)}\left(rR+\Lambda_1(\pi_1(\bar{\pi};t);t')(S-C)\right)dt'\right. \\ &\qquad \left. +S\int_t^\infty e^{-r(t'-t)}\frac{p_1(\pi_1(\bar{\pi};t),t')}{p_1(\pi_1(\bar{\pi};t);t)}(-dq_f^t(t'))\right], \end{aligned} \tag{4.1}$$

where the first integral is the expected present value (EPV) of the payoff stream received until the follower's exit, and the second integral is the EPV of the payoff, when the follower exits prior to the first failure.

Introduce the notation

$$A = \frac{r(R-S)}{C}, \tag{4.2}$$

$$\hat{A}_1(\pi_1(\bar{\pi};t),t) = \max_{t'\geq t}\Lambda_1(\pi_1(\bar{\pi};t),t') = \Lambda_1(\pi_1(\bar{\pi};t),\hat{t}_1(\pi_1(\bar{\pi};t)), \tag{4.3}$$

$$\Phi(A,\bar{\pi},t;T) = C\int_t^T e^{-r(t'-t)}\frac{p_1(\pi_1(\bar{\pi};t);t')}{p_1(\pi_1(\bar{\pi};t);t)}\left(A-\Lambda_1(\pi_1(\bar{\pi};t);t')\right)dt'. \tag{4.4}$$

Lemma 4.1. *The value of the follower, given by equation* (4.1)*, can be equivalently written as*

$$\begin{aligned} F(\bar{\pi};t) &= S \\ &\quad +\sup_{q_f^t} C\int_t^\infty e^{-r(t'-t)}\frac{p_1(\pi_1(\bar{\pi};t);t')}{p_1(\pi_1(\bar{\pi};t);t)}\left(A-\Lambda_1(\pi_1(\bar{\pi};t);t')\right)q_f^t(t')dt'. \end{aligned} \tag{4.5}$$

The first term in representation (4.1) is the value of immediate exit; the second term is the option value of waiting. See Section 9.3 for the proof.

Let $\tau > t$ denote the time of the first failure of the project if the quality is bad. Since experimentation is not profitable after the first failure, we have $\hat{q}^t_f(t') = 0$ for all $t' > \tau$. In all the theorems below, the optimal strategies $\hat{q}^t_f(t')$ are conditioned on $t' \le \tau$. For the brevity of exposition, we omit multiplication of the strategies by the indicator function $\mathbb{1}_{t' \le \tau}$.

Theorem 4.2. *If $A \ge \hat{A}_1(\pi_1(\bar{\pi}; t))$, the only optimal strategy of the follower is*

$$\hat{q}^t_f(t') = 1, \ \forall\, t' > t, \tag{4.6}$$

and

$$F(\bar{\pi}; t) = S + \Phi(\bar{\pi}, t; +\infty). \tag{4.7}$$

Proof. Under condition $A \ge \hat{A}_1(\pi_1(\bar{\pi}; t))$, the integrand on the RHS of (4.5) is non-negative, and positive in a neighborhood of $+\infty$. Hence, the integral is maximized with the choice (4.6). □

Lemma 4.3. *Let $A < \hat{A}_1(\pi_1(\bar{\pi}; t))$. Then (a) the equation*

$$A - \Lambda_1(\pi_1(\bar{\pi}; t); t') = 0 \tag{4.8}$$

has exactly two solutions $0 < t^*_1(A, \pi_1(\bar{\pi}; t)) < t_{*1}(A, \pi_1(\bar{\pi}; t))$.

(b) $t^*_1(A, \pi_1(\bar{\pi}; t))$ is the local maximum of $\Phi(A, \bar{\pi}, t; \cdot)$; $t_{*1}(A, \pi_1(\bar{\pi}; t))$ is the local minimum of $\Phi(A, \bar{\pi}, t; \cdot)$.

Proof. (a) By Lemma 2.3, $\Lambda_1(\pi_1(\bar{\pi}; t), t')$ is hump-shaped. Therefore, if $A < \hat{A}_1(\pi_1(\bar{\pi}; t))$, equation (4.8) has exactly two solutions:

$$0 < t^*_1(A, \pi_1(\bar{\pi}; t) < \hat{t}_1(\pi_1(\bar{\pi}; t)) < t_{*1}(A, \pi_1(\bar{\pi}; t)).$$

(b) If $0 < t^*_1(A, \pi_1(\bar{\pi}; t)) < t_{*1}(A, \pi_1(\bar{\pi}; t))$ are solutions to (4.8), then it is easy to see that the LHS in (4.8) is positive if $t' < t^*_1$ or $t' > t_{*1}$; and it is negative if $t^*_1 < t' < t_{*1}$. Since $\Phi(A, \bar{\pi}, t, \cdot)$ is increasing (respectively, decreasing) iff the LHS in (4.8) is positive (respectively, negative), (b) follows. □

Lemma 4.4. *For any $t \ge 0$, there exists a unique $A^*_1(\pi_1(\bar{\pi}; t)) \in \Big(0, \hat{A}_1(\pi_1(\bar{\pi}; t))\Big)$ s.t.*

$$\Phi(A, \bar{\pi}, t; +\infty) \le 0 \ \Leftrightarrow A \le A^*_1(\pi_1(\bar{\pi}; t)).$$

Proof. Fix $(\bar{\pi}; t)$, and suppress the dependence on $(\bar{\pi}; t)$ in the notation $\hat{A}_1$, $t^*_1(A)$, A^*_1. For $A < \hat{A}_1$, consider

$$\Phi(A, \bar{\pi}, t^*_1(A); +\infty) = C \int_{t^*_1(A)}^{\infty} e^{-rt'} \frac{p_1(\pi_1(\bar{\pi}; t); t')}{p_1(\pi_1(\bar{\pi}; t); t)} \left(A - \Lambda_1(\pi_1(\bar{\pi}; t); t')\right) dt'. \tag{4.9}$$

The integrand in (4.9) increases in A. Furthermore, while A remains below $\hat{A}_1$, the integrand is negative in a right neighborhood of $t^*_1(A)$, and $t^*_1(A)$ moves to the right as A increases. Hence, the integral in (4.9) is increasing in A. As a function of A, the integral is positive at $\hat{A}_1$; by continuity, it is also positive in a left neighborhood of $\hat{A}_1$. In the limit $A \to +0$, the integral becomes negative. Hence, there exists $A_1^* \in (0, \hat{A}_1)$ s.t. the integral in (4.9) is negative for any $A < A_1^*$, and positive for any $A \in (A_1^*, \hat{A}_1)$. By monotonicity of $\Phi(A, \bar{\pi}, t_1^*(A); +\infty)$, this A_1^* is unique. □

Theorem 4.5. *Let $A < \hat{A}_1(\pi_1(\bar{\pi}; t))$. Then*

(1) if $A > A_1^*(\pi_1(\bar{\pi}; t))$, the follower never exits before the first failure, and the only optimal strategy is (4.6)*;*

(2) if $A < A_1^*(\pi_1(\bar{\pi}; t))$ and $t \le t_1^*(A, \pi_1(\bar{\pi}; t))$, the follower exits at $t_1^*(A, \pi_1(\bar{\pi}; t))$. The only optimal strategy is

$$\hat{q}_f^t(t') = \begin{cases} 1, & t' \in (t, t_1^*(A, \pi_1(\bar{\pi}; t))], \\ 0, & t' > t_1^*(A, \pi_1(\bar{\pi}; t)); \end{cases} \tag{4.10}$$

(3) if $A = A_1^*(\pi_1(\bar{\pi}; t))$ and $t \le t_1^*(A, \pi_1(\bar{\pi}; t))$, then, for any $\bar{q} \in [0, 1]$, the strategy

$$\hat{q}_f^t(t') = \begin{cases} 1, & t' \in (t, t_1^*(A, \pi_1(\bar{\pi}; t))], \\ \bar{q}, & t' > t_1^*(A, \pi_1(\bar{\pi}; t)), \end{cases} \tag{4.11}$$

is optimal, and any optimal strategy is of the form (4.11)*;*

(4) if $A < A_1^*(\pi_1(\bar{\pi}; t))$ and $t > t_1^*(A, \pi_1(\bar{\pi}; t))$, then the only optimal strategy is

$$\hat{q}_f^t(t') = 0, \ t' > t; \tag{4.12}$$

(5) if $A = A_1^*(\pi_1(\bar{\pi}; t))$ and $t > t_1^*(A, \pi_1(\bar{\pi}; t))$, then, for any $\bar{q} \in [0, 1]$, the strategy

$$\hat{q}_f^t(t') = \bar{q}, \ \forall\, t' > t, \tag{4.13}$$

is optimal, and any optimal strategy is of the form (4.13)*.*

Obviously, cases (3) and (5) are non-generic, while the rest of the cases in Theorem 4.5 are generic ones. All statements of the theorem are immediate from the following Lemma.

Lemma 4.6. *Let the following conditions hold*

(i) F_1 is a piece-wise continuous function and there exit $B > 0$ and $r > 0$ s.t. $|F_1(t)| \le Be^{-rt}$ $\forall\, t \ge 0$.

(ii) Function F_2 defined by

$$F_2(T) = \int_t^T F_1(t')dt'$$

has a finite number $t \le t_1 < t_2 < \cdots t_n \le +\infty$ of points of the global maximum.

(iii) $q : [t, +\infty) \to [0, 1]$ is a non-decreasing LCRL function with the trivial singular continuous component, s.t. $q(t) = 1$;

Then the problem

$$V = \sup_q \int_t^{+\infty} F_1(t')q(t')dt'$$

where the supremum is taken over the class of function satisfying (iii), has solutions of the form

$$\hat{q}_f^t(t') = \begin{cases} 1, & t' \in [t, t_1], \\ \bar{q}_j, & t' \in (t_j, t_{j+1}], j = 1, 2, \dots, n-1, \\ 0, & t' > t_n, \end{cases} \tag{4.14}$$

where $1 \geq \bar{q}_1 \geq \bar{q}_2 \geq \cdots \geq \bar{q}_n \geq 0$*, and any optimal solution is of the form* (4.14).

Proof. Integrating by parts, we obtain

$$\begin{aligned} V &= -F_2(t) + F_2(+\infty)q(+\infty) + \int_t^{+\infty} F_2(t')(-dq(t')) \\ &= F_2(+\infty)q(+\infty) + \int_t^{+\infty} F_2(t')(-dq(t')). \end{aligned}$$

and the statement follows. □

To summarize, there exists the critical value A_1^* is s.t. for all $A < A_1^*$, $t_1^*(A)$ is the global maximum of $\Phi(A, \bar{\pi}, t; \cdot)$. If $A > A_1^*$, the global maximum of $\Phi(A, \bar{\pi}, t; \cdot)$ is at $T = +\infty$. If $A = A_1^*$, $\Phi(A, \bar{\pi}, t; \cdot)$ (which is a non-generic case) has two maxima - $T = t_1^*(A)$ and $T = +\infty$. The statements of Theorems 4.2 and 4.5 imply that, in a generic case, the follower may find it optimal to exit at the same time as the leader, never to exit before the first failure, or exit some time after the leader's exit unless the first failure happens earlier. Let $t_f = t_f(A, \bar{\pi}; t)$ denote the optimal stopping time of the follower, which may be t, $t_1^*(A, \pi_s(\bar{\pi}, t))$, or $+\infty$. Then, in a generic case, we can write the follower's optimal strategy as

$$\hat{q}_f^t(t') = \begin{cases} 1, & t' \in (t, t_f], \\ 0, & t' > t_f; \end{cases}$$

and the follower's value as

$$F(\bar{\pi}; t) = S + \Phi(A, \bar{\pi}, t; t_f). \tag{4.15}$$

4.2. **Value of player i when both players are active.** Consider a subgame starting at $t \geq 0$ after a history such that none of the players has yet acted. Consider the value function of player i in such a subgame. We have

$$\begin{aligned} V_i(\bar{\pi};t,q_i^t,q_j^t) &= \int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}G_i(\bar{\pi};t')q_i^t(t')q_j^t(t')dt' \\ &+ \int_{\{t'\geq t \ | \ \Delta q_i^t(t')=0\}}^\infty e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi},t)}F_i(\bar{\pi};t')q_i^t(t')(-dq_j^t(t')) \\ &+ \int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}Sq_j^t(t')(-dq_i^t(t')), \end{aligned} \tag{4.16}$$

where, as before, $G_i(\bar{\pi};t')$ is the expected payoff flow that player i gets when both players are experimenting. Assuming that, if the project is bad, the players are equally likely to incur a costly failure, we can write

$$G_i(\bar{\pi};t') = rR + \Lambda_2(\bar{\pi};t')(S-0.5C).$$

Lemma 4.7. *We have*

$$\begin{aligned} V_i(\bar{\pi};t,q_i^t,q_j^t) &= S \\ &+ \int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar{\pi},t')}{p_2(\bar{\pi},t)}\big[C\left(A-0.5\Lambda_2(\bar{\pi};t')\right)q_j^t(t') \\ &\qquad -\Phi(A,\bar{\pi},t';t_f(A,\bar{\pi},t'))(q_j^t(t'))'\big]\, q_i^t(t')dt \\ &+ \sum_{\substack{t'\geq t:\\ \Delta q_i(t')=0\\ \Delta q_j(t')\neq 0}} e^{-r(t'-t)}\frac{p_2(\bar{\pi},t')}{p_2(\bar{\pi},t)}\Phi(A,\bar{\pi},t';t_f(A,\bar{\pi},t'))q_i^t(t')(-\Delta q_j(t')) \\ &- \sum_{\substack{t'\geq t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}} Se^{-r(t'-t)}\frac{p_2(\bar{\pi},t')}{p_2(\bar{\pi},t)}\left(q_i^t(t')+\Delta q_i^t(t')\right)(-\Delta q_j(t')). \end{aligned} \tag{4.17}$$

See Section 9.4 for the proof.

5. Equilibria

5.1. SPE, where both players stay until the first failure.

Let

$$\hat{A}_2 = \hat{A}_2(\bar{\pi}) = 0.5\max_{t\geq 0}\Lambda_2(\bar{\pi};t) = 0.5\Lambda_2(\bar{\pi};\hat{t}_2(\bar{\pi})), \tag{5.1}$$

and

$$\Psi(A,\bar{\pi},t;T) = C\int_t^T e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}\left(A-0.5\Lambda_2(\bar{\pi};t')\right)dt'. \tag{5.2}$$

Let $\tau > t$ denote the time of the first failure of one of the projects if the quality is bad. Since experimentation is not profitable after the first failure, we have $\hat{q}_i^t(t') = \hat{q}_j^t(t') =$

0 for all $t' > \tau$. In all the theorems below, the optimal strategies $\hat{q}_i^t(t')$ ($i \in \{1,2\}$) are conditioned on $t' \le \tau$. For the brevity of exposition we omit multiplication of strategies by the indicator function $\mathbb{1}_{t'\le\tau}$.

Theorem 5.1. *If $A \ge \hat{A}_2(\bar{\pi})$, then for any $t \ge 0$, the following profile is a SPE in the subgame starting at t: for $i, j \in \{1,2\}$, $i \neq j$:*

$$\hat{q}_i^t(t') = \hat{q}_j^t(t') = 1, \quad t' > t, \tag{5.3}$$

and

$$V_i(\bar{\pi}; t, \hat{q}_i^t, \hat{q}_j^t) = V_j(\bar{\pi}; t, \hat{q}_i^t, \hat{q}_j^t) = S + \Psi(\kappa, A, \bar{\pi}, t; +\infty). \tag{5.4}$$

Proof. Notice that the proposed strategies are consistent. Consider a subgame that starts at $t \ge 0$ after such a history that no player has acted as yet. Let player j follow the prescribed strategy (5.3). Then $\Delta q_j^t(t') = 0$ and $(q_j^t(t'))' = 0$ for all $0 \le t \le t'$. Player i chooses the best response $q_i^t(t')$ which solves the following problem

$$\sup_{q_i^t(t')} C \int_t^\infty e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} \left(A - 0.5\Lambda_2(\bar{\pi}; t')\right) q_i^t(t') dt'. \tag{5.5}$$

Let

$$\psi(A, \bar{\pi}, t') := A - 0.5\Lambda_2(\bar{\pi}; t') \tag{5.6}$$

Since $A \ge \hat{A}_d(\bar{\pi})$, $\psi(A, \bar{\pi}, t') \ge 0$ for all $t' \ge t$. Hence the best response of player i to q_j^t given by (5.3) is to play $\hat{q}^t(t') = 1$ for all $t' \ge t$. Hence $(\hat{q}_i^t, \hat{q}_j^t)$ given by (5.3) is a SPE. □

Theorem 5.1 states that if $A = r(R-S)/C$ is sufficiently large, then in a SPE, the players stop simultaneously at the moment of the first failure if the project is bad or never if the project is good.

Lemma 5.2. *Let $A < \hat{A}_2(\bar{\pi})$, then*

a) the equation

$$A - 0.5\Lambda_2(\bar{\pi}; t) = 0 \tag{5.7}$$

has two solutions $t_2^*(A, \bar{\pi}) < t_{*2}(A, \bar{\pi})$ s.t. $t_2^*(A, \bar{\pi})$ is the local maximum, and $t_{*2}(A, \bar{\pi})$ is the local minimum of $\Psi(A, \bar{\pi}, t; \cdot)$.

b) $t_2^*(A, \bar{\pi}) = t_1^*(A, \pi_1(\bar{\pi}; t_2^*))$.

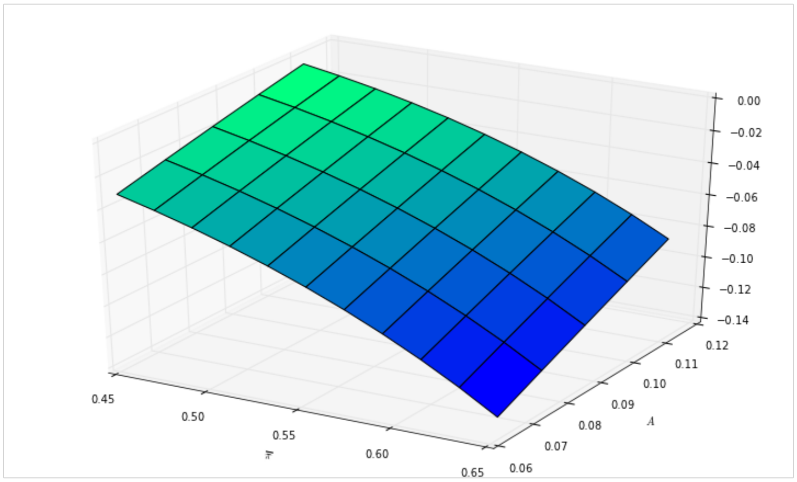
c) If $\Psi(A, \bar{\pi}, t_2^*(A, \bar{\pi}); +\infty) \le 0$, then $\Phi(A, \bar{\pi}, t_2^*(A, \bar{\pi}); +\infty) < 0$.

See Section 9.5 for the proof.

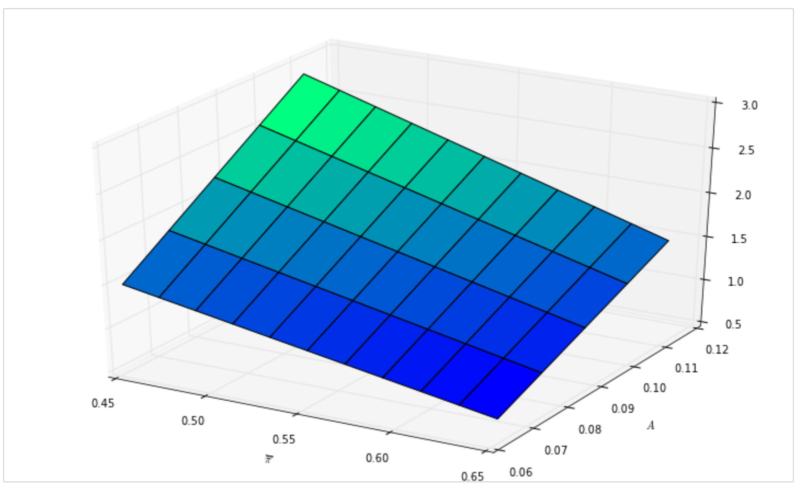
Remark 5.3. Lemma 5.2 implies that

(i) In any subgame that starts at time $t < t_2^* = t_2^*(A, \bar{\pi})$ (we suppress the dependence of t_2^* on $(A, \bar{\pi})$ in order to simplify the notation) after a history such that none of the players has yet acted, none of the players will find it optimal to exit before t_2^*, because $\psi(A, \bar{\pi}, t)$ given by (5.6) is positive for all $t \le t' < t_2^*$. To understand why, recall that

$$C\psi(A, \bar{\pi}, t') = r(R-S) - 0.5\Lambda_2(\bar{\pi}; t')C$$



(A) Erlang-2 model: $\lambda = 1$



(B) Exponential model: $\lambda_e = 0.5$

FIGURE 2. Option value of waiting when two players are active. Parameters: $r = 0.02$, $C = 1$.

is the net expected marginal benefit of each player, when two players keep experimenting. It is not optimal to stop while the net marginal benefit is positive.

(ii) If one of the players (the leader) decides to exit at time $t_2^* = t_2^*(A, \bar{\pi})$, the other player (the follower) will find it optimal to exit together with the leader at $t_2^*(A, \bar{\pi})$ because $\Psi(A, \bar{\pi}, t_2^*(A, \bar{\pi}); +\infty) \leq 0$ implies $\Phi(A, \bar{\pi}, t_2^*(A, \bar{\pi}); +\infty) < 0$.

(iii) Notice that for the same range of pairs $(A, \bar{\pi})$ where players in Erlang-2 model stop in finite time (option value of waiting is negative at $t_2^*(A, \bar{\pi})$, players in the corresponding exponential model never stop before the first failure happens because the option value of waiting is positive at any $0 \leq t \leq \tau$ as Fig. 2 illustrates. We scale the time so that $\lambda = 1$ in the Erlang-2 model, then the expected time of arrival of bad news is 2. We consider the exponential model with the same expected time of arrival of the first piece of new, which implies that in the exponential model, the rate of arrival is $\lambda_e = 0.5$.

We have the following analog of Lemma 4.4.

Lemma 5.4. *There exists a unique $A_2^* = A_2^*(\bar\pi) \in (0, \hat A_2(\bar\pi))$ s.t.*

$$\Psi(A_2^*, \bar\pi, t_2^*; +\infty) = 0, \tag{5.8}$$

and $\Psi(A, \bar\pi, t_2^*; +\infty) > 0 (= 0, < 0)$ if $A > A_2^* (= A_2^*, < A_2^*)$.

Proof follows the proof of Lemma 4.4 line by line.

Theorem 5.5. *Let $A \ge A_2^*(\bar\pi)$, then for $t \ge 0$, $(\hat q_i^t, \hat q_j^t)$ given by (5.3) is a SPE, and the payoffs are given by (5.4).*

Proof. If $A > A_2^* = A_2^*(\bar\pi)$, the function $\Psi(A, \bar\pi, t, T)$ is maximized at $T = +\infty$. Hence if player j plays $\hat q_j^t$, the best response of player i is to play $\hat q_i^t$ and vice versa. By definition, $\Psi(A_2^*, \bar\pi, t_2^*; +\infty) = 0$, hence for any $t > t_2^*$, $\Psi(A_2^*, \bar\pi, t; +\infty) > 0$, hence it is never optimal to exit before the first failure happens, hence $(\hat q_i^t, \hat q_j^t)$ given by (5.3) is a SPE. If $t \le t_2^*$, it is never optimal to exit earlier than at t_2^*, because $\Psi(A, \bar\pi, t, T)$ is increasing in T if $T \in (t, t_2^*]$. Since $\Psi(A_2^*, \bar\pi, t_2^*; +\infty) = 0$, the players are indifferent between exiting at t_2^* and staying until the first failure happens. □

5.2. Symmetric SPE, where players stop before the first failure.

Theorem 5.6. *Let $A < A_2^*(\kappa, \bar\pi)$. Consider a subgame that starts at $t \le t_2^*(A, \bar\pi)$ after a history such that none of the players has yet acted Then, there exists a symmetric SPE given by the following pair of simple consistent strategies:*

$$\hat q_i^t(t') = \hat q_j^t(t') = \begin{cases} 1, & t' \le t_2^*(A, \bar\pi) \\ 0, & t' \ge t_2^*(A, \bar\pi). \end{cases} \tag{5.9}$$

The players payoffs are

$$V_i(\bar\pi, t; \hat q_i^t, \hat q_j^t) = V_j(\bar\pi, t; \hat q_i^t, \hat q_j^t) = S + \Psi(A, \bar\pi, t; t_2^*). \tag{5.10}$$

See Fig. 3 for illustration.

Proof. Since it is not optimal to stop earlier than at $t_2^*(A, \bar\pi)$, and $t_2^*(A, \bar\pi) = t_1(A, \pi_s(\bar\pi; t_2^*))$, simultaneous stopping is an equilibrium on the strength of Lemma 5.2. See Remark 5.3 (ii) for explanation.

Theorem 5.6 specifies strategies on the equilibrium path. Next, we specify strategies off the equilibrium path. Let $A < A_2^*(\bar\pi)$, and let $\hat T_2 = \hat T_2(A, \bar\pi) > t_2^*(A, \bar\pi)$ be the solution to

$$\Psi(A, \bar\pi, \hat T_2, +\infty) = 0.$$

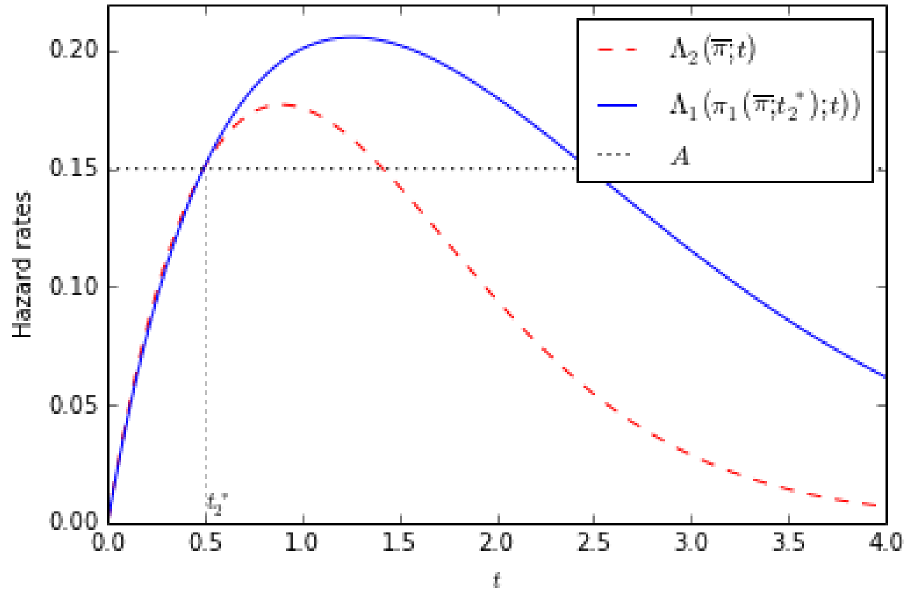


FIGURE 3. Hazards rates and optimal stopping times in the Erlang-2 bad news model. Parameters: $\lambda = 1$, $A = 0.15$, $r = 0.02$, $\bar{\pi} = 0.5$

Then, for $t < \hat{T}_2$, $\Psi(A, \bar{\pi}, t, +\infty) < 0$, and for $t > \hat{T}_2$, $\Psi(A, \bar{\pi}, t, +\infty) > 0$. Hence, for any subgame that starts at $t \in (t_2^*, \hat{T}_2)$ after a history such that none of the players has yet acted, the optimal strategies are $\hat{q}_i^t(t') = \hat{q}_j^t(t') = 0$. For any subgame that starts at $t \geq \hat{T}_2$ after a history such that none of the players has yet acted, the optimal strategies are $\hat{q}_i^t(t') = \hat{q}_j^t(t') = 1$. Due to the existence of the non-empty interval $(t_2^*, \hat{T}_2)$, strategies (5.9) are consistent. Hence, simultaneous stopping is a SPE. □

Remark 5.7. 1. The results of Lemma 5.2 and Theorem 5.6 can be extended to the case of $n > 2$ players. Let t_n^* denote the smallest solution to

$$A - \frac{1}{n}\Lambda_n(\bar{\pi}; t) = 0.$$

In this case, all the players exit at t_n^*.

2. It is easy to see that simultaneous exit at t_n^* is efficient, because stopping at t_n^* is also a solution to the social planner's problem.

6. CONCLUSIVE BREAKTHROUGHS

6.1. **The setup.** The stylized model presented here is suitable to describe a grant competition or innovation contest, where the winner (if any) takes the prize, and the remaining players get the outside option. We will indicate later how the results may change if the winner takes the major prize, but remaining players are also rewarded with the reward being higher than the value of the outside option. Assume that $n \geq 2$ symmetric players experiment with technologies of unknown quality. The quality of a project depends on the state of nature $\theta \in \{0, 1\}$. If $\theta = 1$, the project is good, which means that it generates positive revenues (breakthroughs). If $\theta = 0$, the project is bad, which means that the player experimenting with this project will never be able to generate positive revenues. Experimentation is costly, and the stream of experimentation costs is rC for each player, independent of the quality of the project. The initial common prior assigns probability $\bar{\pi} \in (0, 1)$ to $\theta = 1$.

Let the quality of the project be the same for both players, so that if one of the players observes a breakthrough, both players know that their technologies are good. For simplicity, assume that if the project if good, active players stop experimentation after the first success had been observed. Further assume that, if the project is good, the player who is first to succeed gets $R > 0$, and the other player(s) get $S \in [0, R)$. Thus, there is an advantage to generating the success first (though, in this stylized model, this advantage is independent of the players' actions), which can be interpreted as an opportunity to file a patent. If a player exits before the first observation of a success, then the player gets S. So, S is the value of the outside option.

We use the same primitives as described in Section 2 and same strategies as in Section 3.

6.2. **Value functions and equilibrium.** Consider the game that starts at $t = 0$. Unlike in the bad news model, starting experimentation at $t = 0$ may not be optimal, because the rate of arrival of good news (profitable payoff) is close to zero in a right neighborhood of zero. We leave for the future study strategic entry decisions of the players in this experimentation game, where we expect free riding and encouragement effects to be present. In this paper, we study only strategic exit decisions of active players, who start experimentation at the same time, due to, for example deadline specified by designers of the grant competition. To this end, we assume that

$$\sup_{T>0} \left(\int_0^T e^{-rt} p_n(\bar{\pi}; t) \left((R + (n-1)S) \frac{\Lambda_n(\bar{\pi}; t)}{n} - rC \right) dt + e^{-rT} p_n(\bar{\pi}; T) S \right) > S. \tag{6.1}$$

Then it is optimal for n players to experiment until the time $\hat{t}_n(\bar{\pi})$ when the rate of arrival reaches its maximal value

$$\hat{t}_n(\bar{\pi}) = \arg\max_{t \geq 0} \Lambda_n(\bar{\pi}; t),$$

and for some time after that. As the rate of arrival $\Lambda_n(\bar{\pi}; t)$ starts decreasing, it may become optimal for one or more players to quit. We will show that the stopping rules in pure strategies are of the threshold type - the players quit when the corresponding rates of arrival reach a certain threshold from above.

Lemma 6.1. *There exists $n^* \geq 1$ s.t. inequality* (6.1) *is satisfied for all $n < n^*$ and is not satisfied for $n \geq n^*$.*

See Section 9.6 for the proof. Lemma 6.1 demonstrates the crowding out effect of experimentation. In the corresponding exponential model such effect never takes place - if experimentation is optimal for one player, it is optimal for any number of players.

From now on, we will focus on the case $n^* > 2$ and consider the strategic experimentation game between two players. The results admit straightforward generalization to the case $n > 2$.

Let $G_i(\bar{\pi}; t) = 0.5\Lambda_2(\bar{\pi}; t)(R + S) - rC$ denote the instantaneous expected payoff flow of player i if none of the players stopped until time $t > 0$. Let $F_i(\bar{\pi}; t)$ denote the expected value of player i if player j stopped at time t, and player i did not. Given the strategy profile (q_i, q_j), the value of player i in the game that starts at $t = 0$ is

$$\begin{aligned} V_i(\bar{\pi}; q_i, q_j) \quad &= \quad \int_0^\infty e^{-rt} p_2(\bar{\pi}; t) G_i(\bar{\pi}; t) q_i(t) q_j(t) dt \qquad (6.2) \\ &+ \int_{\{t \geq 0 \ | \ \Delta q_i(t) = 0\}} e^{-rt} p_2(\bar{\pi}, t) F_i(\bar{\pi}; t) q_i(t)(-dq_j(t)) \\ &+ \int_0^\infty e^{-rt} p_2(\bar{\pi}; t) S q_j(t)(-dq_i(t)). \end{aligned}$$

Later we will show that $F_i(\bar{\pi}; \cdot)$ is continuous and has the finite limit as $t \to \infty$, hence, $V(\bar{\pi}; q_i, q_j)$ is well-defined and finite. Note that the second integral in (6.2) takes into account jumps in q_j only, and last integral takes into account jumps in q_i only as well as simultaneous jumps in q_i and q_j. Once one of the players has quitted experimentation, the other player faces a non-strategic stopping problem, which can be easily solved. Thus, when considering subgame perfect equilibria, we will first examine subgames when one of the players has stopped, and then move to subgames where neither player has quitted as yet. To simplify the notation, we suppress the dependence of value functions on the other player's strategy. Since the players are symmetric, we also drop the subscripts identifying the players.

6.3. **Follower's problem.** Consider a subgame that starts after the history such that only one of the players has stopped. Suppose, this happened at time t. Then the remaining player (the follower) chooses a strategy q_f^t satisfying the conditions of Definition 3.2, which solves the following problem:

$$\begin{aligned} F(\bar{\pi}; t) \quad &= \quad \sup_{q_f^t} \Bigg[\int_t^\infty e^{-r(t'-t)} q_f^t(t') \frac{p_1(\pi_1(\bar{\pi}; t); t')}{p_1(\pi_1(\bar{\pi}; t); t)} \left(\Lambda_1(\pi_1(\bar{\pi}; t); t') R - rC\right) dt' \\ &+ S \int_t^\infty e^{-r(t'-t)} \frac{p_1(\pi_1(\bar{\pi}; t); t')}{p_1(\pi_1(\bar{\pi}; t); t)} (-dq_f^t(t')) \Bigg], \qquad (6.3) \end{aligned}$$

where the first integral is the expected present value (EPV) of the payoff while the follower is active, and the second integral is the EPV of the payoff when the follower exits prior to the first failure.

Let

$$A = \frac{r(C+S)}{R-S} > 0. \tag{6.4}$$

Similarly to Lemma 4.1, we obtain

Lemma 6.2. *The value of the follower, given by equation* (6.3), *can be equivalently written as*

$$\begin{aligned} F(\bar{\pi}, t) &= S \\ &+ (R-S)\sup_{q_f^t} \int_t^{\infty} e^{-r(t'-t)} \frac{p_1(\pi_1(\bar{\pi};t);t')}{p_1(\pi_1(\bar{\pi};t);t)} \left(\Lambda_1(\pi_1(\bar{\pi};t);t') - A\right) q_f^t(t')dt'. \end{aligned} \tag{6.5}$$

The proof is analogous to the proof of Lemma 4.1. The first term in representation (6.2) is the value of immediate exit; the second term is the option value of waiting. Since $\Lambda_1(\pi_1(\bar{\pi};t);t') \to 0$ as $t' \to +\infty$, the follower exits either instantly if either $\Lambda_1(\pi_1(\bar{\pi};t),t') \le A$ and t is to the right from the point $\hat{t}_1 = \arg\max_{t' \ge t} \Lambda_1(\pi_1(\bar{\pi};t);t')$; or $t < \hat{t}_1$ and $\max_{T \ge t} \Phi(A, \bar{\pi}, t; T) \le 0$, where

$$\Phi(A, \bar{\pi}, t; T) = (R-S) \int_t^T e^{-r(t'-t)} \frac{p_1(\pi_1(\bar{\pi};t);t')}{p_1(\pi_1(\bar{\pi};t);t)} \left(\Lambda_1(\pi_1(\bar{\pi};t);t') - A\right) dt'$$

(in the case of equality, it is also optimal to wait until the local maximizer T is achieved). Otherwise, the follower exits at time $T_1 = T_1(A, \pi_1(\bar{\pi};t)) < +\infty$, $T_1 > t$, which is the largest solution of the following equation

$$\Lambda_1(\pi_1(\bar{\pi};t), T) = A. \tag{6.6}$$

Thus, the follower's exit time $t_f = t_f(A, \pi_1(\bar{\pi};t))$ can equal to t or $T_1(A, \pi_1(\bar{\pi};t))$. In a generic case, we can write the follower's optimal strategy as

$$\hat{q}_f^t(t') = \begin{cases} 1, & \forall\, t \le t' \le t_f(A, \pi_1(\bar{\pi};t)), \\ 0, & \forall\, t' > t_f; \end{cases}$$

and the follower's value as

$$F(\bar{\pi}, t) = S + \Phi(A, \bar{\pi}, t; t_f(A, \pi_1(\bar{\pi};t))). \tag{6.7}$$

6.4. **Value of player i when two players are active.** Consider a subgame starting at $t \ge 0$ after a history such that none of the players has yet acted. Consider the value function of player i in such a subgame. We have

$$\begin{aligned} V_i(\bar{\pi}, t; q_i^t, q_j^t) &= \int_t^\infty e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} G_i(\bar{\pi}; t') q_i^t(t') q_j^t(t') dt' \\ &+ \int_{\{t' \geq t \ \mid \ \Delta q_i^t(t')=0\}}^\infty e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} F_i(\bar{\pi}; t') q_i^t(t')(-dq_j^t(t')) \\ &+ \int_t^\infty e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} S q_j^t(t')(-dq_i^t(t')). \end{aligned} \tag{6.8}$$

The following result can be proved in the same manner as Lemma 4.7.

Lemma 6.3. *We have*

$$\begin{aligned} V_i(\bar{\pi}, t; q_i^t, q_j^t) &= S \\ &+ \int_t^\infty e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} \left[(R-S)\left(0.5\Lambda_2(\bar{\pi}; t') - A\right) q_j^t(t') \right. \\ &\qquad \left. - \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t'))(q_j^t(t'))' \right] q_i^t(t') dt' \\ &+ \sum_{\substack{t' \geq t: \\ \Delta q_i(t') = 0 \\ \Delta q_j(t') \neq 0}} e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t')) q_i^t(t')(-\Delta q_j(t')) \\ &- \sum_{\substack{t' \geq t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} S e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} \left(q_i^t(t') + \Delta q_i^t(t') \right) (-\Delta q_j(t')). \end{aligned} \tag{6.9}$$

Since $\Lambda_2(\bar{\pi}; t) \to 0$ as $t \to +\infty$, it is non-optimal to experiment jointly either after time t or after time $T = t_2(A, \bar{\pi}) < +\infty$, which is the largest solution of the following equation

$$0.5\Lambda_2(\bar{\pi}; T) = A. \tag{6.10}$$

Conditions for the instantaneous exit can be derived in the same way as in the case of failures. We consider the more interesting case, when $t < t_2(A, \bar{\pi})$, and, for all $t < t_2(A, \bar{\pi})$, it is non-optimal to stop the joint experimentation.

Lemma 6.4. *Let $t < t_2(A, \bar{\pi})$. Let $t_1(A, \pi_1(\bar{\pi}; t_2))$ denote the largest solution to* (6.6) *for $t = t_2(A, \bar{\pi})$. Then $t_1(A, \pi_1(\bar{\pi}; t_2)) = t_2(A, \bar{\pi})$, hence both players exit at $t_2(A, \bar{\pi})$.*

Proof. By Lemma 2.2,

$$\Lambda_1(\pi_1(\bar{\pi}; t_2); t_2) = 0.5\Lambda_2(\bar{\pi}; t_2) = A,$$

therefore, $t_2 = t_2(A, \bar{\pi})$ is a solution to (6.6). To prove that it is the largest solution to (6.6), we use the following argument. Since $t_2 = t_2(A, \bar{\pi})$ is the largest solution to

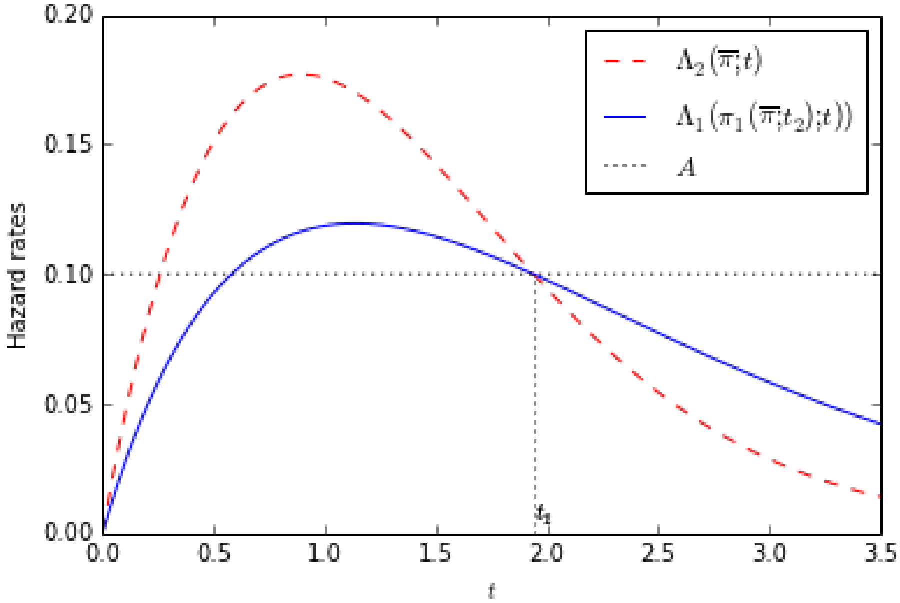


FIGURE 4. Hazards rates and optimal stopping times in the Erlang-2 good news model. Parameters: $\lambda = 1$, $A = 0.1$, $r = 0.02$, $\bar{\pi} = 0.5$

(6.10), $t_2(A, \bar{\pi}) > \hat{t}_2(\bar{\pi})$. Therefore, on the strength of (9.2),

$$\partial_t \left(\frac{-1}{\Lambda_1(1; t_2)} \right) > 2(1 - \pi_2(\bar{\pi}, t_2)) > 1 - \pi_2(\bar{\pi}, t_2) = 1 - \pi_1(\pi_1(\bar{\pi}, t_2), t_2).$$

Hence $t_2(A, \bar{\pi}) > \hat{t}_1(\pi_1(\bar{\pi}, t_2))$, hence $t_2 = t_2(A, \bar{\pi})$ is the largest solution to (6.6). □

Theorem 6.5. *Let $t < t_2 = t_2(A, \bar{\pi})$. There exists a unique SPE defined by*

$$\hat{q}_i^t(t') = \hat{q}_j^t(t') = \begin{cases} 1, & t' \in (t, t_2], \\ 0, & t' > t_2. \end{cases} \tag{6.11}$$

The players' payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = S + \Psi(A, \bar{\pi}, t; t_2), \tag{6.12}$$

where

$$\Psi(A, \bar{\pi}, t; T) = (R - S) \int_t^T e^{-r(t'-t)} \frac{p_2(\bar{\pi}; t')}{p_2(\bar{\pi}; t)} \left(0.5\Lambda_2(\bar{\pi}; t') - A\right) dt'. \tag{6.13}$$

See Fig. 4 for illustration.

Proof. Since it is not optimal to stop earlier than at $t_2(A, \bar{\pi})$, and $t_2(A, \bar{\pi}) = t_1(A, \pi_1(\bar{\pi}; t_2))$, simultaneous stopping is an equilibrium. □

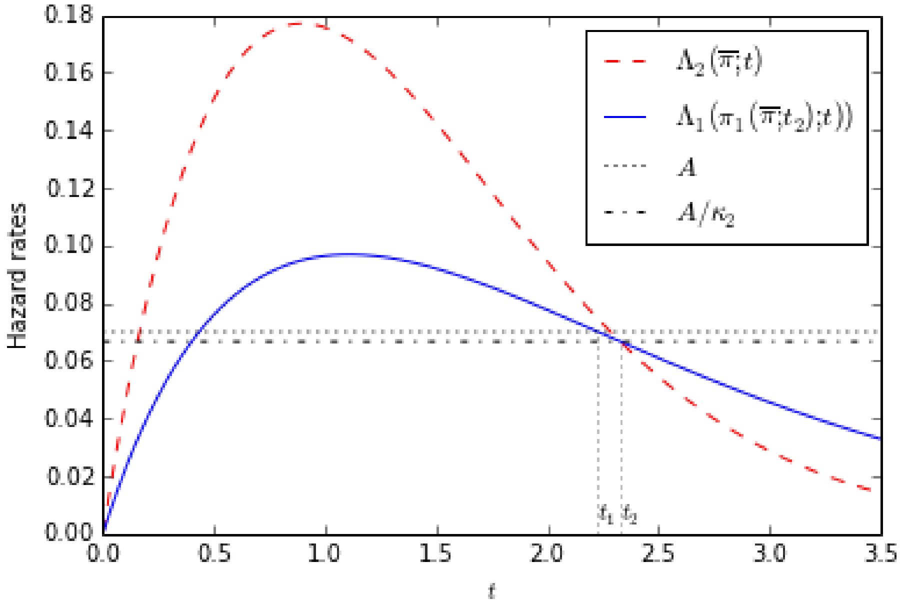


FIGURE 5. Hazards rates and optimal stopping times in the Erlang-2 good news model with payoff externalities. Parameters: $n = 2$, $\lambda = 1$, $A = 0.07$, $r = 0.02$, $\bar{\pi} = 0.5$, $\kappa_2 = 1.05$.

Remark 6.6. 1. The result of Theorem 6.5 can be extended to the case of $n > 2$ players if the following inequalities hold for each $n > 1$:

$$\int_0^{t_n} e^{-rt} p_n(\bar{\pi}; t) \left(A - \frac{1}{n} \Lambda_n(\bar{\pi}; t) \right) dt > 0,$$

where t_n is the largest solution to

$$\frac{1}{n} \Lambda_n(\bar{\pi}; t) - A = 0. \tag{6.14}$$

In this case, all the players exit at t_n.

2. It is easy to see that simultaneous exit at t_n is efficient, because stopping at t_n is also a solution to the social planner's problem.

We see that the critical value of the hazard rates at exit is independent of the number of players. In order the critical values of hazard rates depended on n, one may introduce payoff externalities in addition to information externalities. For example, suppose, for example, that the player, who gets the first breakthrough gets the prize $R > S$, and the rest of the players get payoffs $S < P < R$. Then each player's expected payoff in case of arrival of good news is $(R + (n-1)P)/n$. Let, as before,

$A = r(C+S)/(R-S)$, and let

$$\kappa_n = 1 + \frac{(n-1)(P-S)}{R-S}.$$

Then the exit threshold for n players is the largest solution to

$$\frac{1}{n}\Lambda_n(\bar{\pi};t) - \frac{A}{\kappa_n} = 0, \tag{6.15}$$

provided

$$\int_0^{t_n} p_n(\bar{\pi};t)\left(\frac{1}{n}\Lambda_n(\bar{\pi};t) - \frac{A}{\kappa_n}\right)dt > 0. \tag{6.16}$$

The last inequality is similar to (6.1) (see the proof of Lemma 6.1 in Section 9.6 to understand why). Therefore, there exists n^* s.t. (6.16) holds iff $n < n^*$. If inequality (6.16) is satisfied for some $1 < n < n^*$, then players exit simultaneously at t_n, because

$$t_1(t_2,\dots,t_n) < \cdots t_{n-1}(t_n) < t_n,$$

where $t_{n-1}(t_n)$ is the largest solution to

$$\frac{1}{n-1}\Lambda_{n-1}(\pi_1(\bar{\pi};t_n);t) - \frac{A}{\kappa_{n-1}} = 0,$$

and so on. Fig. 5 and Fig. 6 illustrate the above for the case of two and three players, respectively. Since $\kappa_n > 1$, the largest solution to (6.15) is bigger than the largest solution to (6.14), so the players will experiment longer. Moreover, inequality (6.16) will be satisfied for a larger number of players than (6.1), so the crowding out effect will be weaker. Thus, if a designer of a grant competition has in mind her own optimal length of experimentation, longer experimentation can be achieved by a proper reward scheme design. This conclusion is true, provided that costs spent on experimentation by every player are observable. If costs are private information, then it is necessary to study incentives to free ride on other players' experimentation as, e.g., in Halac et.al [14].

7. Extensions and generalization

7.1. **Hump-shaped distributions.** We derived results in the paper under a standing assumption that $\Lambda_1(1;0) = 0$. This assumption is convenient to study bad news model, because it is always optimal to start experimentation since the rate of arrival of news is close to zero in a right neighborhood of zero. At the same time, such standing assumption may be rather restrictive. In this Section, we generalize our model further and provide classification of one-humped bandits. As in Section 2, the primitive is the rate of arrival of news when the risky project can generate news for sure. Also, as before, assume that $n \geq 1$ players experiment with a project of unknown quality, and $\bar{\pi} \in [0,1]$ is the common prior which is the probability of the event that the project generates news.

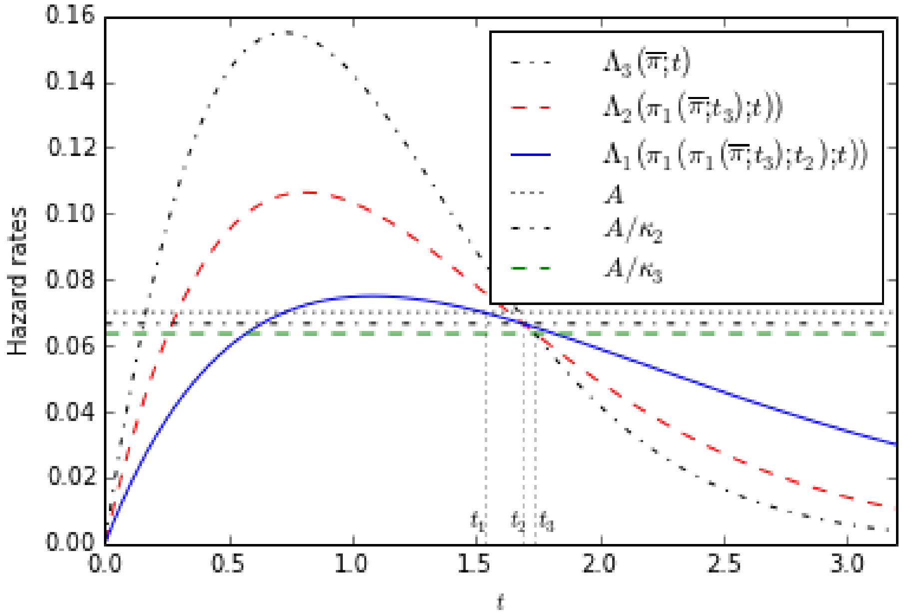


FIGURE 6. Hazards rates and optimal stopping times in the Erlang-2 good news model with payoff externalities. Parameters: $n = 3$, $\lambda = 1$, $A = 0.07$, $r = 0.02$, $\bar{\pi} = 0.5$, $\kappa_2 = 1.05$, $\kappa_3 = 1.1$.

Lemma 7.1. *Let*

(i) $\Lambda_1(1;t)$ *be of the class* C^2 *and increasing in* t*;*

(ii)

$$\partial_t^2 \left(\frac{1}{\Lambda_1(1;t)} \right) > 0, \ \forall t > 0; \tag{7.1}$$

(iii)

$$\lim_{t \downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)} \right) > n(1 - \bar{\pi}); \tag{7.2}$$

(iv) and

$$\lim_{t \to \infty} \partial_t \left(\frac{-1}{\Lambda_1(1;t)} \right) < 1.$$

Then

(a) as a function of t, $\Lambda_n(\bar{\pi};t) = n\Lambda_1(1;t)\pi_n(\bar{\pi};t)$ *has the global maximum* $\hat{t}_n = \hat{t}_n(\bar{\pi})$*;*
(b) If (7.1) holds but (7.2) fails, then $\Lambda_n(\bar{\pi};t)$ *is a decreasing function on* $\mathbb{R}_+$*.*

If

$$1 - \bar{\pi} < \lim_{t \downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)} \right) < \hat{n},$$

for some $1 < \hat{n} \le n$, then the qualitative behavior of the arrival rate $\Lambda_n(\bar{\pi}; t)$ depends on n. Namely, for $n \ge \hat{n}$,, the arrival rate $\Lambda_n(\bar{\pi}; t)$ is a decreasing function, and the model is qualitatively the same as the exponential bandits model.

7.2. **Classification of one-humped bandits.** Next, we characterize different possible types of one-humped bandits. We start with the following preliminary remarks.

- If $\lim_{t\downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)}\right) = +\infty$, then (7.2) holds for any n.
- If $0 < \lim_{t\downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)}\right) < +\infty$, then $1/\Lambda_1(1;t)$ is bounded as $t \to 0$, and $\Lambda_1(1;0) > 0$.
- It is possible that $\Lambda_1(1;0) > 0$, (7.1) holds but $\lim_{t\downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)}\right) = +\infty$.
- If $\partial_t \Lambda_1(1;0)$ exists and (7.1) holds, then $\Lambda_1(1;+0) > 0$ iff

$$0 < \lim_{t\downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)}\right) < +\infty.$$

Definition 7.2. Let $\partial_t^2 \left(\frac{1}{\Lambda_1(1;t)}\right) > 0, \ \forall t > 0$.

We call the bandit model defined by $\Lambda_1(1;t)$ a one-humped model of Type I, II and III if the corresponding condition below holds

I. $\Lambda_1(1;0) = 0$, and $\partial_t \Lambda_1(1;0)$ exists, and it is finite;
II. $\Lambda_1(1;0) > 0$ and $\partial_t \Lambda_1(1;0)$ exists, and it is finite;
III. $\Lambda_1(1;0) > 0$ and $\lim_{t\downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1;t)}\right) = +\infty$.

7.3. **One-humped bandits of Types II and III.** Properties specified for Type I bandits in the previous Sections hold, and equilibria of the same types are possible.

Depending on the parameters, the usual encouragement effect can be observed (as in exponential bandit models).

An additional effect and type of equilibria (if r is sufficiently large): discouragement (crowding out) effect: $\exists m \ge 1$ s.t.

(1) if $n < m$ players are in the game at time 0, they will find it optimal to start experimenting with the bad news technology;
(2) $n \ge m$ players will not start experimenting unless $n - m$ of them exit instantly.

7.4. **Multi-humped bandits.** We call the model a multi-humped model, if $\Lambda_n(\bar{\pi}; t)$ has more than one point of local maximum. Examples

(a) the environment with some seasonality;
(b) if business cycle effects are taken into account;
(c) endogenous multi-humped bandits.

7.5. **Endogenous multi-humped bandits.** Assume that the players plan to enter the game with breakdowns at times $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n < t_{n+1} := +\infty$; this can be an equilibrium outcome if, for example, players are asymmetric.

Then the rate of arrival $\Lambda_k(\bar{\pi}; t)$, which the k players that are in the game face, is defined as follows. For $k = 1, 2, \ldots, n$ and $t \in [t_k, t_{k+1})$,

$$\Lambda_k(\bar{\pi}; t) = \sum_{j=1}^{k} \Lambda_j(\bar{\pi}; t - t_j).$$

Clearly, more than one hump is possible, and if the underlying one-humped bandit is of Type II or III, then Λ_k exhibits jumps.

8. Conclusion

We suggested a new model for strategic experimentation, where good or bad news arrive at random times which are modeled as jump time of a time-inhomogeneous Poisson process. This models are almost as tractable as exponential bandits models and can incorporate such realistic features as dependence of the expected rate of news arrival on the time elapsed since the start of an experiment. In this paper, we

- characterized SPE in a model with conclusive failures;
- characterized optimal stopping strategies in terms of critical levels of expected rates of arrival of time-inhomogeneous Poisson processes;
- showed that depending on parameters of the model the following equilibrium outcomes are possible:
  (i) none of the players find it optimal to stop before the first failure;
  (ii) players stop simultaneously before the first piece of bad news arrives;
- similar results hold in a model with breakthroughs;
- the model with breakthroughs demonstrates the crowding out effect, which is absent in exponential models;
- suggested a classification of humped bandits.

In the future, we plan to extend the current model to inconclusive experiments, correlated arms, private payoffs, and other types of humped bandits. We believe that humped shaped are quite promising and can be used not only as a framework in learning and experimentation models, but in other areas of economics or finance. The can also be combined with other sources of uncertainty. For example, credit risk models with hump-shaped hazard rate and evolution of assets modeled as a jump-diffusion process are tractable but more realistic than oversimplified reduced form models of defaultable debt with the exponential hazard rate.

References

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

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

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

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

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

[6] 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).

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

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

[9] 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)

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

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

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

[13] P.K. Dutta, and A. Rustichini "A Theory of Stopping Time Games With Applications to Product Innovations and Asset Sales, *Economic Theory*, 3 (1993), 743-763.

[14] M. Halac, N. Kartik, and Q. Liu, "Contests for experimentation," to appear in *Journal of Political Economy*, 125:5 (2017), 1523-1569.

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

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

[17] 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

[18] 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

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

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

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

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

[23] N. Klein, and S. Rady, "Negatively correlated bandits," *Review of Economic Studies*, 78:2 (2008), 693-732.

[24] U. Khan, and M.B. Stinchcombe, "The Virtues of Hesitation: Optimal Timing in a Non-stationary World," *American Economic Review*, 105:3 (2015), 1147-1176.

[25] R. Laraki, E. Solan, and N. Vieille, "Continuous-time Games of Timing," *Journal of Economic Theory*, 120 (2005), 206–238.

[26] M. Ludkovski, and R. Sircar, "Exploration and Exhaustibility in Dynamic Cournot Games," *European Journal of Applied Mathematics*, 23 (2012), 343-372.
[27] P. Murto, and J. Välimäki, "Delay and information aggregation in stopping games with private information," *Journal of Economic Theory*, 148 (2013), 2404-2435.
[28] K. Roberts, and M.L. Weitzman, "Funding criteria for research, development, and exploration projects," *Econometrica*, 49:5 (1981), 1261-1288.
[29] D. Rosenberg, A. Salomon, and N. Vieille, "On games of strategic experimentation," *Games and Economic Behavior*, 82 (2013), 31-51.
[30] M. Rothschild, "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 9:2 (1974), 185-202.
[31] M.L. Weitzman, "Optimal search for the best alternative," *Econometrica*, 47:3 (1979), 641-654.

9. Appendix

9.1. **Proof of Lemma 2.1.** (i) It suffices to rewrite (2.6) as

$$\pi_n(\bar\pi;t) = 1 - \frac{1-\bar\pi}{p_n(\bar\pi;t)},$$

and notice that $p_n(\bar\pi;t)$ is decreasing in n because $p_1(1;t) < 1$ for all $t > 0$.
(ii) Let $\hat\pi$ be a solution to

$$\pi_n(\bar\pi, t') = \pi_{n-1}(\hat\pi, t')\,. \tag{9.1}$$

We write (9.1) as

$$\frac{\bar\pi p_1(1;t')^n}{1-\bar\pi+\bar\pi p_1(1;t')^n} = \frac{\hat\pi p_1(1;t')^{n-1}}{1-\hat\pi+\hat\pi p_1(1;t')^{n-1}}.$$

Equivalently,

$$\hat\pi(1-\bar\pi)p_1(1;t') = (1-\hat\pi)\bar\pi,$$

whence

$$\hat\pi = \frac{\bar\pi p_1(1;t')}{1-\bar\pi+\bar\pi p_1(1;t')} = \pi_1(\bar\pi;t),$$

and (2.8) follows.
(iii) Follows from strict monotonicity of $\pi_n(\bar\pi;t)$ and $\pi_{n-1}(\pi_1(\bar\pi;t'),t)$ in t.
(iv) Follows from (iii) and the fact that

$$\pi_n(\bar\pi,0) = \bar\pi > \pi_1(\bar\pi;t') = \pi_{n-1}(\pi_1(\bar\pi;t'),0).$$

9.2. **Proof of Lemma 2.2.** (a) Intersection at $t=0$ follows from (2.10) and property (i) $\Lambda_1(1;0) = 0$. By Lemma 2.1, and equation (2.10), the only positive point of intersection is t'.
(b) It follows from (2.10) and the equality in (2.7) that

$$\begin{aligned}\frac{1}{n}\partial_t\Lambda_n(\bar\pi;t) &= \partial_t\Lambda_1(1;t) + \Lambda_1(1;t)\partial_t\pi_n(\bar\pi;t)\\ &= \pi_n(\bar\pi;t)\left[\partial_t\Lambda_1(1;t) - n\Lambda_1(1;t)^2(1-\pi_n(\bar\pi;t))\right].\end{aligned} \tag{9.2}$$

Therefore,

$$\frac{1}{n}\partial_t\Lambda_n(\bar\pi;+0) = \bar\pi\partial_t\Lambda_1(1;+0) > \pi_1(\bar\pi;t')\partial_t\Lambda_1(1;+0) = \frac{1}{n-1}\partial_t\Lambda_{n-1}(\pi_1(\bar\pi;t'),0).$$

By continuity,

$$\frac{1}{n}\partial_t\Lambda_n(\bar\pi;t) > \frac{1}{n-1}\partial_t\Lambda_{n-1}(\pi_1(\bar\pi;t'),t)$$

in a right neighborhood of zero. Hence,

$$\frac{1}{n}\Lambda_n(\bar\pi;t) > \frac{1}{n-1}\Lambda_{n-1}(\pi_1(\bar\pi;t'),t)$$

in a right neighborhood of zero. Since there is only one positive intersection of the hazard rates, (2.12) follows.

9.3. **Proof of Lemma 4.1.** We start with the following result.

Lemma 9.1. *Let the following conditions hold:*

(i) $f:[0,+\infty)\to\mathbb{R}$ *be continuous and differentiable;*
(ii) there exists $B>0$ *and* $r>0$ *s.t.* $|f(t)|\le Be^{-rt}\ \forall\, t\ge 0$;
(iii) q *satisfies the condtions of Definition 3.1 and the singular continuous component of the Lebesgue decomposition of* q *is trivial.*

Then

$$\int_t^\infty f(t')(-dq(t')) = f(t)q(t) + \int_t^\infty f'(t')q(t')dt'. \tag{9.3}$$

Proof. By definition,

$$\int_t^\infty f(t')(-dq(t')) = -\int_t^\infty f(t')q'(t')dt' + \sum_{\substack{t'\ge t:\\ \Delta q(t')\neq 0}} f(t')(-\Delta q(t')),$$

and

$$\begin{aligned}
-f(t)q(t) &= \int_t^\infty d\left(f(t')q(t')\right)\\
&= \int_t^\infty \left(f(t')q(t')\right)' dt' - \sum_{\substack{t'\ge t:\\ \Delta q(t')\neq 0}} f(t')(-\Delta q(t'))\\
&= \int_t^\infty f'(t')q(t')dt' + \int_t^\infty f(t')q'(t')dt' - \sum_{\substack{t'\ge t:\\ \Delta q(t')\neq 0}} f(t')(-\Delta q(t'))\\
&= \int_t^\infty f'(t')q(t')dt' - \int_t^\infty f(t')(-dq(t')).
\end{aligned}$$

Equation (9.3) follows. □

Using Lemma 9.1, we rewrite the second integral in (4.1) as

$$\begin{aligned}
&S\int_t^\infty e^{-r(t'-t)}\frac{p_1(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t))}(-dq^t(t'))\\
&= -Se^{-r(t'-t)}\frac{p_1(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t))}q^t(t')\Big|_t^\infty + S\int_t^\infty \frac{d}{dt'}\left(e^{-r(t'-t)}\frac{p_1(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t))}\right)q^t(t')dt'\\
&= S + S\int_t^\infty e^{-r(t'-t)}\frac{p_1(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t))}\left(-r+\frac{p_1'(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t')}\right)q^t(t')dt'\\
&= S + S\int_t^\infty e^{-r(t'-t)}\frac{p_1(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t))}\left(-r-\Lambda_1(\pi_1(\bar\pi;t);t')\right)q^t(t')dt'.
\end{aligned}$$

Substitute the last expression for the the second integral in (4.1), then

$$F(\bar\pi;t) = S + \sup_{q^t}\int_t^\infty e^{-r(t'-t)}\frac{p_1(\pi_1(\bar\pi;t);t')}{p_1(\pi_1(\bar\pi;t);t))}\left(r(R-S)-\Lambda_1(\pi_1(\bar\pi;t),t')C\right)q^t(t')dt'.$$

Dividing and multiplying the integrand by C and using (4.2), we arrive at (4.1).

9.4. **Proof of Lemma 4.7.** We start with the following result.

Lemma 9.2. *Let the following conditions hold:*

(i) $f:[0,+\infty)\to\mathbb{R}$ *is continuous and differentiable;*
(ii) there exists $B>0$ *and* $r>0$ *s.t.* $|f(t)|\le Be^{-rt}\ \forall\, t\ge 0$;
(iii) q_i, q_j *satisfy the conditions of of Definition 3.1, and the singular continuous components of the Lebesgue decomposition of* q_i, q_j *are trivial.*

Then

$$\begin{aligned}
&\int_t^\infty f(t')q_j(t')(-dq_i(t')) \qquad (9.4)\\
=\;& -f(t)q_i(t)q_j(t) + \int_t^\infty f'(t')q_i(t')q_j(t')dt'\\
&-\int_t^\infty f(t')q_i(t')(-dq_j(t')) - \sum_{\substack{t'\ge t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}} f(t')(-\Delta q_i(t')\Delta q_j(t')).
\end{aligned}$$

Proof. Applying Definition 3.4, we obtain

$$
\begin{aligned}
-f(t)q_i(t)q_j(t) &= \int_t^\infty d\left(f(t')q_i(t')q_j(t')\right) \\
&= \int_t^\infty \left(f(t')q_i(t')q_j(t')\right)' dt' - \sum_{\substack{t' \geq t: \\ \Delta(q_i(t')q_j(t')) \neq 0}} f(t')(-\Delta(q_i(t')q_j(t'))) \\
&= \int_t^\infty f'(t')q_i(t')q_j(t')dt' + \int_t^\infty f(t')q_i'(t')q_j(t')dt' + \int_t^\infty f(t')q_j'(t')q_i(t')dt' \\
&\quad - \sum_{\substack{t' \geq t: \\ \Delta q_i(t') \neq 0}} f(t')q_j(t')(-\Delta q_i(t')) - \sum_{\substack{t' \geq t: \\ \Delta q_j(t') \neq 0}} f(t')q_i(t')(-\Delta q_j(t')) \\
&\quad - \sum_{\substack{t' \geq t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} f(t')(-\Delta q_i(t')\Delta q_j(t')) \\
&= \int_t^\infty f'(t')q_i(t')q_j(t')dt' - \int_t^\infty f(t')q_j(t')(-dq_i(t'))
\end{aligned}
$$

$$
- \int_t^\infty f(t')q_i(t')(-dq_j(t')) - \sum_{\substack{t' \geq t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} f(t')(-\Delta q_i(t')\Delta q_j(t')).
$$

Equation (9.4) follows. ∎

Now we can prove Lemma 4.7, in several steps.

Step 1. Rewrite the second integral on the RHS of (4.16) as

$$\int_{\{t'\geq t \ | \ \Delta dq_i^t(t')=0\}}^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}F_i(\bar{\pi};t')q_i^t(t')(-dq_j^t(t'))$$
$$= \int_t^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}F_i(\bar{\pi};t')q_i^t(t')(-dq_j^t(t'))$$
$$- \sum_{\substack{t'\geq t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}F_i(\bar{\pi};t')q_i^t(t')(-\Delta q_j(t'))$$
$$= -\int_t^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}F_i(\bar{\pi};t')q_i^t(t')(q_j^t(t'))'dt' \tag{9.5}$$
$$+ \sum_{\substack{t'\geq t:\\ \Delta q_i(t')= 0\\ \Delta q_j(t')\neq 0}} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}F_i(\bar{\pi};t')q_i^t(t')(-\Delta q_j(t')).$$

Substitute for $F_i(\bar{\pi};t')$ on the RHS of equation (5.9). Then (9.5) becomes

$$\int_{\{t'\geq t \ | \ \Delta dq_i^t(t')=0\}}^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}F_i(\bar{\pi};t')q_i^t(t')(-dq_j^t(t'))$$
$$= -\int_t^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}(S+\Phi(A,\bar{\pi},t';t_f(A,\bar{\pi},t')))q_i^t(t')(q_j^t(t'))'dt' \tag{9.6}$$
$$+ \sum_{\substack{t'\geq t:\\ \Delta q_i(t')= 0\\ \Delta q_j(t')\neq 0}} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}(S+\Phi(A,\bar{\pi},t';t_f(A,\bar{\pi},t')))q_i^t(t')(-\Delta q_j(t')).$$

Step 2. Using Lemma 9.2 write the last integral in (4.16) as

$$\int_t^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}Sq_j^t(t')(-dq_i^t(t')) \tag{9.7}$$
$$= S\left[-e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}q_j^t(t')q_i^t(t')\Big|_t^{\infty} + \int_t^{\infty}\frac{d}{dt}\left(e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}\right)q_i^t(t')q_j^t(t')dt'\right.$$
$$\left.-\int_t^{\infty} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}q_i^t(t')(-dq_j^t(t')) - \sum_{\substack{t'\geq t:\\ \Delta q_i^t(t')\neq 0\\ \Delta q_j^t(t')\neq 0}} e^{-r(t'-t)}\frac{p_2(\bar{\pi};t')}{p_2(\bar{\pi};t)}(-\Delta q_i^t(t')\Delta q_j^t(t'))\right].$$

Using the fact that the first term in the brackets is equal to one,, we rewrite (9.7) as

$$\int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}Sq^t_j(t')(-dq^t_i(t')) = S \tag{9.8}$$
$$+S\int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}\left(-r-\Lambda_2(\bar\pi;t')\right)q^t_i(t')q^t_j(t')dt'$$
$$+S\int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)})q^t_i(t')(q^t_j(t'))'dt' - \sum_{\substack{t'\ge t:\\ \Delta q_j(t')\neq 0}} Se^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}q^t_i(t')(-\Delta q_j(t')).$$
$$-\sum_{\substack{t'\ge t:\\ \Delta q^t_i(t')\neq 0\\ \Delta q^t_j(t')\neq 0}} Se^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}(-\Delta q^t_i(t')\Delta q^t_j(t'))\Big].$$

Adding (9.6) and (9.8) to the first integral in (4.16), we obtain

$$\begin{aligned}
V_i(\bar\pi,t;q^t_i,q^t_j) &= S+\\
&\times\int_t^\infty e^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}\left[\left(r(R-S)-0.5\Lambda_2(\bar\pi;t')C\right)q^t_j(t')\right.\\
&\qquad\left.-\Phi(A,\bar\pi,t';t_f(A,\bar\pi,t'))(q^t_j(t'))'\right]q^t_i(t')dt'\\
&+\sum_{\substack{t'\ge t:\\ \Delta q_i(t')=0\\ \Delta q_j(t')\neq 0}} e^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}\Phi(A,\bar\pi,t';t_f(A,\bar\pi,t'))q^t_i(t')(-\Delta q_j(t'))\\
&-\sum_{\substack{t'\ge t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}} Se^{-r(t'-t)}\frac{p_2(\bar\pi;t')}{p_2(\bar\pi;t)}\left(q^t_i(t')+\Delta q^t_i(t')\right)(-\Delta q_j(t')).
\end{aligned}$$

It remains to use the notation $A = r(R-S)/C$ to arrive at (4.17).

9.5. **Proof of Lemma 5.2.** To prove (a), one can repeat the proof of Lemma 4.3 line by line.

(b) On the strength of Lemma 2.2, functions $\Lambda_1(\pi_1(\bar\pi,t^*_2(\bar\pi,A));t)$ and $0.5\Lambda_2(\bar\pi;t)$ have only one positive point of intersection at $t = t^*_2(\bar\pi,A)$, and

$$0.5\Lambda_2(\bar\pi;t) > \Lambda_1(\pi_1(\bar\pi,t^*_2(\bar\pi,A));t),\ \ t < t^*_2(\bar\pi,A).$$

Since $t^*_2(\bar\pi,A)$ is a solution to (5.7), $\Lambda_1(\pi_1(\bar\pi,t^*_2(\bar\pi,A));t^*_2(\bar\pi,A)) = A$, and (b) follows.

(c) It suffices to show that for all $t \geq t_2^*$,

$$\frac{p_1(\pi_1(\bar{\pi};t_2^*);t)}{p_1(\pi_1(\bar{\pi};t_2^*);t_2^*)}\left(A-\Lambda_1(\pi_1(\bar{\pi};t_2^*);t)\right) < \frac{p_2(\bar{\pi};t)}{p_2(\bar{\pi};t_2^*)}\left(A-0.5\Lambda_2(\bar{\pi};t)\right). \tag{9.9}$$

By (2.5),

$$p_1(\pi_1(\bar{\pi};t_2^*);t) = 1-\pi_1(\bar{\pi};t_2^*)+\pi_1(\bar{\pi};t_2^*)p_1(1;t).$$

Substituting (2.6) for $\pi_1(\bar{\pi};t_2^*)$, it is straightforward to show that

$$p_1(\pi_1(\bar{\pi};t_2^*);t) = \frac{1-\bar{\pi}+\bar{\pi}p_1(1;t_2^*)p_1(1;t)}{p_1(\bar{\pi};t_2^*)}, \tag{9.10}$$

in particular,

$$p_1(\pi_1(\bar{\pi};t_2^*);t_2^*) = \frac{1-\bar{\pi}+\bar{\pi}p_1(1;t_2^*)^2}{p_1(\bar{\pi};t_2^*)} = \frac{p_2(\bar{\pi};t_2^*)}{p_1(\bar{\pi};t_2^*)}.$$

Therefore, (9.9) is equivalent to

$$\left(1-\bar{\pi}+\bar{\pi}p_1(1;t_2^*)p_1(1;t)\right)\left(A-\Lambda_1(\pi_1(\bar{\pi};t_2^*);t)\right) < p_2(\bar{\pi};t)\left(A-0.5\Lambda_2(\bar{\pi};t)\right). \tag{9.11}$$

Using (2.2) for the arrival rate $\Lambda_n(\bar{\pi};t)$, it is straightforward to show that

$$\begin{aligned} 0.5\Lambda_2(\bar{\pi};t)p_2(\bar{\pi};t) &= \Lambda_1(1;t)\bar{\pi}p_1(1;t)^2; \\ \Lambda_1((\pi_1(\bar{\pi};t_2^*);t))\left(1-\bar{\pi}+\bar{\pi}p_1(1;t_2^*)p_1(1;t)\right) &= \Lambda_1(1;t)\bar{\pi}p_1(1;t_2^*)p_1(1;t). \end{aligned}$$

Therefore, (9.11) is equivalent to

$$Ap_1(1;t)(p_1(1;t_2^*)-p_1(1;t)) < \Lambda_1(1;t)p_1(1;t)(p_1(1;t_2^*)-p_1(1;t)).$$

Since $p_1(1;t) > 0$ and $p_1(1;t_2^*) - p_1(1;t)$ (because $p_1(1;t)$ is decreasing in t), we have that (9.11) is equivalent to

$$A < L_1(1;t),$$

which is true for because $t > t_2^*$, $L_1(1;t)$ is increasing in t, and, by (5.7),

$$L_1(1;t_2^*) = \frac{A}{\pi_2(\bar{\pi};t_2^*)} > A.$$

Hence (c) follows.

9.6. **Proof of Lemma 6.1.** Rewrite inequality (6.1) using the following steps. (i) Use integration by parts to calculate the integral

$$\begin{aligned} \int_0^T re^{-rt}p_n(\bar{\pi};t)dt &= 1-e^{-rT}p_n(\bar{\pi};T)+\int_0^T e^{-rt}\partial_t p_n(\bar{\pi};t)dt \\ &= 1-e^{-rT}p_n(\bar{\pi};T)-\int_0^T e^{-rt}p_n(\bar{\pi};t)\Lambda_n(\bar{\pi};t)dt. \end{aligned}$$

The second equality follows from the fact that

$$\begin{aligned}\partial_t p_n(\bar\pi;t) &= \bar\pi\partial_t p_1^n(1,t) = -n\Lambda_1(1,t)\bar\pi p_1^n(1,t)\\ &= -\frac{n\Lambda_1(1;t)\bar\pi p_1^n(1,t)}{p_n(\bar\pi;t)}\cdot p_n(\bar\pi;t)\\ &= -n\Lambda_1(1;t)\pi_n(\bar\pi;t)p_n(\bar\pi;t) = \Lambda_n(\bar\pi;t)p_n(\bar\pi;t).\end{aligned}$$

Hence, we can write

$$Se^{-rT}p_n(\bar\pi;T) - S = -S\int_0^T e^{-rt}p_n(\bar\pi;t)\left(r+\Lambda_n(\bar\pi;t)\right)dt,$$

therefore (6.1) is equivalent to

$$\sup_{T>0}\left(\int_0^T e^{-rt}p_n(\bar\pi;t)\left((R-S)\frac{\Lambda_n(\bar\pi;t)}{n} - r(C+S)\right)dt\right) > 0. \tag{9.12}$$

(ii) Using the definition (6.4) of A, rewrite (9.12) as

$$\sup_{T>0}\left(\int_0^T e^{-rt}p_n(\bar\pi;t)\left(\frac{\Lambda_n(\bar\pi;t)}{n} - A\right)dt\right) > 0. \tag{9.13}$$

Let

$$\hat A_n = \hat A_n(\bar\pi) = \frac{1}{n}\max_{t\ge 0}\Lambda_n(\bar\pi;t).$$

By (2.11), for every $\bar\pi$ and every $t>0$, $\Lambda_n(\bar\pi;t)/n$ is decreasing in n, therefore, for every $\bar\pi$, $\hat A_n(\bar\pi)$ is decreasing in n. Hence, for every pair $(\bar\pi, A)$ there exists n' such that $\hat A_{n'}(\bar\pi)\le A$, and $\hat A_{n'-1}(\bar\pi) > A$. Hence, inequality (9.13) does not hold for all $n\ge n'$. Consider $n\in[1,n')$. Then equation

$$\frac{\Lambda_n(\bar\pi;t)}{n} = A$$

has two solutions $t_n^*(A,\bar\pi) < t_{*n}(A,\bar\pi)$, so that

$$\Lambda_n(\bar\pi;t)/n - A > 0 \;\Leftrightarrow\; t\in(t_n^*(A,\bar\pi), t_{*n}(A,\bar\pi)).$$

Hence

$$t_{*n}(A,\bar\pi) = \arg\max_T \int_0^T e^{-rt}p_n(\bar\pi;t)\left(\frac{\Lambda_n(\bar\pi;t)}{n} - A\right)dt.$$

Now, we can rewrite (9.13) as

$$\int_0^{t_n^*} e^{-rt}p_n(\bar\pi;t)\left(\frac{\Lambda_n(\bar\pi;t)}{n} - A\right)dt + \int_{t_n^*}^{t_{*n}} e^{-rt}p_n(\bar\pi;t)\left(\frac{\Lambda_n(\bar\pi;t)}{n} - A\right)dt > 0. \tag{9.14}$$

Since $\Lambda_n(\bar\pi;t)/n$ is decreasing in n, t_n^* is increasing, and t_{*n} is decreasing in n. Evidently, the second (positive) integral in (9.14) is decreasing in n, and the absolute value of the first (negative) integral is increasing in n. If (9.14) is not satisfied for $n=1$, then $n^*=1$, and it is not optimal to experiment for any number of players. Let (9.14) be satisfied for $n=1$. If (9.14) holds for any $n=n'-1$, then $n^*=n'$.

Otherwise, there exists a unique $n^* \in (1, n' - 1)$ such that (9.14) holds for $n < n^*$, and does not hold for $n \geq n^*$.