RESEARCH ARTICLE

Super- and submodularity of stopping games with random observations

Svetlana Boyarchenko[1]

Abstract
Models of learning and experimentation based on two-armed Poisson bandits addressed several important aspects related to strategic and motivational learning, but they are not suitable to study effects that accumulate over time. We propose 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. In these models, the experiment is stopped before news is realized whenever the rate of arrival of news reaches a critical level. This leads to longer experimentation times for experiments with possible breakthroughs than for equivalent experiments with failures. We also show that the game with conclusive failures is supermodular, and the game with conclusive breakthroughs is submodular.

Keywords Stopping games · Supermodular games · Time-inhomogeneous Poisson process

✉ Svetlana Boyarchenko
  sboyarch@utexas.edu

[1] The University of Texas at Austin, 2225 Speedway Stop C3100, Austin, TX 78712, USA

1 Introduction

This paper brings together two strands of literature: stopping games in continuous time with random observations and super(sub)modular games. The theory of supermodular games has provided a powerful toolbox for analyzing the consequences of complementarities in economics, including Bayesian, dynamic, and stochastic games (see Vives (2005) for an extensive review of games with complementarities and references therein). While the class of games with strategic complementarities allows for fairly general strategy spaces, including functional spaces, to the best of our knowledge, stopping games have been out of the scope of the theory of supermodular games.

We believe that using the insights from the theory of supermodular games in stopping games can become a promising new approach to stopping games in cases when existence of an equilibrium is questionable. In the current paper, we study supermodularity of strategic exit problems in two models of experimentation with random arrival of observations and unknown rates of arrival.

In many real-life situations, shocks arrive at random and at unknown rates. A pharmaceutical company engaging in drug development knows neither whether the research it engages in will be successful in creating a new drug nor when such success will happen. A bank loan officer does not know whether the borrower will loose her source of income, defaulting on the loan, and is uncertain about the probability with which such an event can occur. Moreover, the arrival rate of shocks in such situations is often increasing in the horizon: the pharmaceutical company can hardly expect the drug research to have immediate success even if the drug is successful for sure; the loan officer would not make the loan if he expected the borrower to default immediately even if the default is imminent. For network goods, such as electronic devices, one can hardly expect high demand right at the time when a network good is introduced. However, if the good is successful, demand may jump up when the "critical mass" of consumers using the same device or application reaches a certain level (see, e.g., Amir and Lazzati (2011) for a standard static model, and Laussel and Resende (2014) for a continuous-time dynamic model, with network externalities).

The main feature of standard models of experimentation based on Poisson (or exponential) bandits is that the rate of arrival of news is constant if the quality of the risky arm is known. We propose a new class of experimentation models which remains nearly as tractable as Poisson bandits models but incorporates both of the features described earlier: unknown quality of the risky arm and, conditional on the quality of the risky arm, an arrival rate of news that is increasing over time. In contrast to exponential bandits models, the models suggested here can be used as a framework for experimentation related to 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 studying network effects; designing grant competitions for long-term projects; or determining optimal terms of politicians. We call this class of models *humped bandits* because the rate of arrival of news is hump-shaped over time: the rate

of arrival is increasing in time from the start of the experiment until it reaches a certain maximal level; after that point, the arrival rate decreases and approaches exponential decay as the experiment "grows older." Intuitively, the longer the experiment lasts without any negative shocks realizing, the more optimistic the experimenter is about the quality of the risky arm. At the same time, since the arrival rate of news conditional on the quality of the risky arm is increasing in the time since the start of the experiment, negative shocks are more likely to happen. Thus, it may be optimal to terminate the experiment before the first negative shock realizes.[1] In such situations, the experiment is terminated while the arrival rate of news is still increasing and, importantly, before any negative news is realized. In particular, in settings with possible failures (e.g., borrower default), the player stops experimenting when the marginal benefit from staying active (e.g., interest rate payments on the loan) equals the expected marginal cost (e.g., losses from a non-performing loan). This paper demonstrates that for a generic set of parameters, there is an equilibrium, where players stop simultaneously before any bad news arrives. In contrast, in the corresponding Poisson bandits models with decreasing arrival rates, the experiment is only terminated after a negative shock occurs.

Experiment termination prior to the realization of negative news is an important realistic prediction of humped bandit models. Some recent high-profile examples of agents acting in line with this prediction include Goldman Sachs' exit from the subprime mortgage market almost a year before the start of the subprime crisis, and Fixed Income Clearing Corporation's (FICC) suspension of its interbank clearing service for general collateral repurchase agreements. In the former case, Goldman Sachs' Chief Financial Officer urged for a more bearish attitude toward the subprime mortgage market as early as December 2006, leading to the structured product division of their mortgage markets group to bet against the subprime mortgage market well in advance of any losses on subprime mortgages being realized.[2] In the latter case, on July 15, 2016, FICC suspended the execution of repurchase agreements involving participants at different clearing banks. In the official filing with the Securities and Exchange Commission (SEC), FICC argued that because the current market design is not amenable to minimizing the amount of intraday credit that FICC requires from the clearing banks to settle interbank transactions, interbank transactions pose an undue operational risk on FICC.[3] Notably, as in the case with Goldman Sachs exiting the subprime mortgage market before the subprime losses were realized, FICC exited the interbank transaction market before any losses from such activity were realized.

Optimal termination prior to the realization of negative news may also have important policy applications. Indeed, the rate of arrival of news, instead of being an increasing function of time, can be also modeled as an increasing function of a different state variable, for example the amount of a fertilizer accumulated in soil or water. On the one hand, the fertilizer can provide benefits to farmers, on the other hand, it can contaminate water and damage ecosystems. The methods provided in the paper would

---

[1] For example, the majority of smokers quit smoking long before being diagnosed with lung cancer. See, e.g., https://www.medscape.com/viewarticle/725138.

[2] See, for example, the Wall Street Journal coverage of the trade in December 2007: https://www.wsj.com/articles/SB119759714037228585.

[3] See SEC filling No. SR-FICC-2016-002.

allow regulators to determine both the optimal rate of application and the optimal amount of the new fertilizer accumulated in soil when further application should stop even though no damage was detected earlier.

Nowadays, the majority of population believes in climate change caused by accumulation of greenhouse gas (GHG). My model can be modified to the case of so-called inconclusive bad news, when players have to learn about the frequency of adverse environmental events whose unknown arrival rate depends on the level of accumulated GHG and decide when it is optimal to invest in abatement technologies. Exponential models will give an incorrect recommendation because they do not take into consideration that hurricanes, flooding, heat waves, etc., may happen more frequently as more GHG emissions are accumulated.

In settings with potential breakthroughs, instead, experimentation does not stop while the rate of arrival of good news is increasing. Intuitively, the longer the experiment lasts without producing a positive outcome, the more pessimistic the experimenter becomes about the quality of the risky arm. At the same time, since the arrival rate of news conditional on the quality of the risky arm is once again increasing in the time since the start of the experiment, positive shocks are more likely to happen. Thus, in settings with potential breakthroughs (e.g., discovery of a new drug), the experiment will get terminated when the marginal cost of the experiment (e.g., R&D expenditure) equals the expected marginal benefit (e.g., profit from selling the drug), and the arrival rate of news is decreasing.[4]

This paper is related to the extensive literature on experimentation based on the so-called two-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. Models of strategic 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, Bergemann and Välimäki (1996, 2000, 2006, 2008), Bolton and Harris (1999), Decamps and Mariotti (2004), Hörner et al. (2014), Keller et al. (2005) and Keller and Rady (2010, 2015). Recent developments include (but are not limited to) correlated risky arms as in Klein and Rady (2008) and Rosenberg et al. (2013), private payoffs as in Heidhues et al. (2015) and Rosenberg et al. (2013), departures from Markovian strategies as in Hörner et al. (2014), strategic costly observations as in Marlats and Ménager (2018), and asymmetric players as in Das et al. (2019). For other developments and an excellent comprehensive review of the literature, see Hörner and Skrzypacz (2016) and references therein.

Poisson bandits models fall into two broad classes: bad news models and good news models. Bad news are costly failures, and good news are profitable successes;

---

[4] Khan and Stinchcombe (2015) 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.

the type of news is known before an experiment starts. The present paper focuses on the case of conclusive failures and successes, so that the first event observed on the risky arm reveals its quality completely, and the experiment is over. It is most closely related to Keller et al. (2005), Keller and Rady (2010, 2015) and Rosenberg et al. (2013). Keller et al. (2005) and Keller and Rady (2010) study good news model, and Keller and Rady (2015) study bad news model; in either model, news arrive at the jump times of independent Poisson processes. Also, Keller et al. (2005) and Keller and Rady (2010, 2015) solve control problems of time allocation between the risky and riskless arms of Poisson bandits., while the present paper considers stopping time problems. Rosenberg et al. (2013) consider an irreversible exit problem in a model with breakthroughs with correlated risky arms both in the case when payoffs are public and private. In contrast, we study a problem of strategic irreversible exit both in good and bad news models, where news arrive at jump times of time-inhomogeneous Poisson processes. Processes that generate news for individual experimenters are independent, but have the same characteristics.

In stopping games, each player chooses a stopping time, and the game ends at the minimum of these stopping times. We specify players' strategies as behavior stopping times as in Laraki et al. (2005). We guess and verify that there is an equilibrium in the class of threshold strategies: for each player there is a time T^* (which depends on actions of the other player, and can be infinite if a player never stops before the first observation arrives) such that the player stays in the game with probability one up to T^* and with probability zero after T^*. In general, threshold-type strategies may be not rich enough to describe equilibria in all stopping games. In particular, if there exists a subset of the state space, called preemption (respectively, attrition) region, where one of the players has the first (respectively, second) mover's advantage but stopping (respectively, staying) together is the worst outcome. In the latter cases, it is necessary to introduce extended mixed strategies, as, for example, in Fudenberg and Tirole (1985), Thijssen et al. (2006, 2012), Boyarchenko and Levendorksiĭ (2014), or Riedel and Steg (2017). However, if the exogenous stochastic process has no jumps in the direction of the preemption or attrition region, then extended mixed strategies are usually needed only off the equilibrium path. For example, Pawlina and Kort (2006) consider only threshold-type strategies, because they do not study subgame perfect equilibria.

The idea of exploiting supermodularity in dynamic games is not new and goes back to Amir (1996a), Curtat (1996) and Vives (2009). In stochastic dynamic games with strategic complementarities, the main focus is on *multi-stage* games with supemodular stage games (see, e.g., Balbus et al. 2014 and references therein). In this paper, instantaneous payoffs of the players are affected by the presence of other players only in terms of rates of information arrival. The threshold-type strategies are formulated in terms of time needed for the rate of arrival to reach a certain critical level. Both in the bad and good news models, the critical level is reached faster when two players are active. At the same time, we show that the stopping game is supermodular in the bad news model, and submodular in the good news model.

The rest of the paper is organized as follows. The primitives of the experimentation model are described in Sect. 2. The stopping game is described in Sect. 3. The latter section also demonstrates that players payoff functions are supermodular in the bad

news model and submodular in the good news model. The detailed analysis of the stopping game in the bad news model and construction of subgame perfect equilibria are in Sect. 4. The stopping game in a good news model is analyzed in Sect. 5. We also study efficiency of equilibria and show that the equilibrium outcome is efficient in the bad news model and inefficient in the good news model. More general intensities of the news arrival are considered in Sect. 6. Section 7 concludes. Technical proofs are relegated to the appendix.

2 Model description

2.1 The setup

Time $t \in \mathbb{R}_+$ is continuous, and the discount rate is $r > 0$. Let two 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). We focus on the case of conclusive experiments, where the first news arrival reveals the quality of the project completely. The quality depends on the state of 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 jump 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 or cause serious side effects. In this paper, we follow the tradition in the bulk of experimentation literature and assume that the type of news is known before the experimentation starts.

Assume that the quality of the projects is the same for both players, so that if one of the players receives a piece of news, the other player concludes that her/his project generates the same type of news for sure. Processes for news arrival are independent of each other, but have the same characteristics. The common prior assigns probability $\pi_0 \in [0, 1]$ 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 the 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 $\pi_0 = 1$. 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 Sect. 6, 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 $\Lambda_1(1; t) = \lambda^2 t/(1+\lambda t)$, where $\lambda > 0$. This hazard rate corresponds to the case when the time of news arrival is an Erlang$(2, \lambda)$ random variable.[5] Note that, in this case, the expected time until the first observation is $2/\lambda$. As $t \to +\infty$, $\Lambda_1(1; t) \to \lambda$.

If $\pi_0 = 1$, the probability of the event that no news is observed before t by a single experimenter is

$$p_1(1; t) = \exp\left(-\int_0^t \Lambda_1(1; s)\mathrm{d}s\right). \tag{2.1}$$

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

If $\pi_0 < 1$, the probability of the event that no news is observed before t by one experimenter is

$$p_1(\pi_0; t) = 1 - \pi_0 + \pi_0 p_1(1; t). \tag{2.2}$$

Obviously, $p_1(\pi_0; 0) = 1$, and $\lim_{t\to\infty} p_1(\pi_0; t) = 1 - \pi_0$. Furthermore, $p_1(\pi_0; t)$ is decreasing in t, because

$$\partial_t p_1(\pi_0; t) = \pi_0 \partial_t p_1(1; t) = -\pi_0 \Lambda_1(1; t) p_1(1; t) < 0 \,\forall t > 0; \tag{2.3}$$

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

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

$$p_2(\pi_0; t) = 1 - \pi_0 + \pi_0 p_1(1; t)^2. \tag{2.4}$$

If no news is observed until time t, when $n \in \{1, 2\}$ players experiment, the posterior probability $\pi_n(\pi_0; t)$ of the event $\theta = 1$ is calculated by Bayes' rule:

$$\pi_n(\pi_0; t) = \frac{\pi_0 p_1(1; t)^n}{p_n(\pi_0; t)}. \tag{2.5}$$

Straightforward differentiation shows that

$$\partial_t \pi_n(\pi_0; t) = -n\Lambda_1(1; t)\pi_n(\pi_0; t)(1 - \pi_n(\pi_0; t)) < 0 \,\forall t > 0, \, n \in \{1, 2\}. \tag{2.6}$$

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 *For any* $t' > 0$,

(i) $\pi_2(\pi_0; t') < \pi_1(\pi_0; t')$;

---

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

(ii) *for any $t' > 0$, there exists a unique $\hat{\pi}$ s.t. $\pi_2(\pi_0; t') = \pi_1(\hat{\pi}, t')$, and*

$$\hat{\pi} = \pi_1(\pi_0; t'); \tag{2.7}$$

(iii)
$$\pi_2(\pi_0; t) < \pi_1(\pi_1(\pi_0; t'), t) \Leftrightarrow t > t'. \tag{2.8}$$

The first statement indicates that any level of beliefs will be reached faster if more players experiment, starting with the same prior. The second statement says that if one of the two players stops experimentation at t', the posterior belief of the remaining player is continuous at t' if π_0 is replaced by $\hat{\pi}$. The last statement indicates that the belief of the remaining player decreases slower than beliefs of two players if none of the players had stopped at t'. See Sect. 8.1 for the proof.

Next, we calculate the rate of arrival of observations at time t when $n \in \{1, 2\}$ players experiment with the projects of unknown quality: $\Lambda_n(\pi_0; t) = n\Lambda_1(1; t)\pi_n(\pi_0; t)$. Since the players are symmetric, each of them is equally likely to receive the first news, therefore the individual hazard rate is

$$\frac{1}{n}\Lambda_n(\pi_0; t) = \Lambda_1(1; t)\pi_n(\pi_0; t), \quad n \in \{1, 2\}. \tag{2.9}$$

Since $\pi_2(\pi_0; t) < \pi_1(\pi_0; t)$ for all $t > 0$,

$$\frac{1}{2}\Lambda_2(\pi_0; t) < \Lambda_1(\pi_0; t) \ \forall \, t > 0. \tag{2.10}$$

The next result describes changes in the hazard rates when the first of two 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) *for any $t' > 0$, functions $0.5\Lambda_2(\pi_0; t)$ and $\Lambda_1(\pi_1(\pi_0; t'); t)$ intersect only at $t = 0$ and $t = t'$;*

(b) $\dfrac{1}{n}\partial_t \Lambda_n(\pi_0; t) = \pi_n(\pi_0; t)\left[\partial_t \Lambda_1(1; t) - n\Lambda_1(1; t)^2(1 - \pi_n(\pi_0; t))\right], \quad n \in \{1, 2\};$

$$\tag{2.11}$$

(c)
$$0.5\Lambda_2(\pi_0; t) < \Lambda_1(\pi_1(\pi_0; t'), t) \Leftrightarrow t > t'. \tag{2.12}$$

See Sect. 8.2 for the proof.

The following result describes the shape of the hazard 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) $\partial_t \Lambda_1(1; 0+)$ *exists and it is finite;*

(iv) $2(\partial_t \Lambda_1(1; t))^2 > \partial_t^2 \Lambda_1(1; t) \cdot \Lambda_1(1; t)$, $\forall t > 0$.

Then as a function of t, $\Lambda_n(\pi_0; t)$ has a unique maximum at $\hat{t}_n = \hat{t}_n(\pi_0)$. Moreover, $\partial_t \Lambda_n(\pi_0; t)$ is positive (respectively, negative) on $(0, \hat{t}_n)$ (respectively, $(\hat{t}_n, \infty)$).

Remark 2.4 Notice that condition (iv) is a very mild restriction on $\Lambda_1(1; t)$. In particular, all concave and linear functions satisfy this condition. All increasing power and exponential functions also satisfy condition (iv).

Proof It suffices to prove that $\partial_t \Lambda_n(1; t)$ is a product of a positive function and a decreasing function which changes sign. Rewrite (2.11) as

$$\frac{1}{n} \partial_t \Lambda_n(\pi_0; t) = \Lambda_1(1; t)^2 \pi_n(\pi_0; t) \left[\frac{\partial_t \Lambda_1(1; t)}{\Lambda_1(1; t)^2} + n\pi_n(\pi_0; t) - n \right], \ n \in \{1, 2\}.$$

The function $f(t) := \partial_t \Lambda_1(1; t)/\Lambda_1(1; t)^2 + n\pi_n(\pi_0; t) - n$ is strictly decreasing because assumption (iv) holds and π_n is decreasing in t. By assumptions (i) and (iv), $\lim_{t \to 0+} f(t) = +\infty$; and $\lim_{t \to \infty} f(t) = -n(1 - \pi_0) < 0$. $\qquad \square$

Lemma 2.3 states that the hazard rate $\Lambda_n(\pi_0; t)$ $(n \in \{1, 2\})$ is hump-shaped. In the exponential model, the corresponding hazard rate $n\lambda\pi_n(\pi_0; t)$ is proportional to the beliefs. Since the beliefs are decreasing in time, the hazard rate is also decreasing.

3 Stopping game

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 Sect. 2. All payoffs, function $\Lambda_1(1; t)$, and the players' actions are public information. 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 sth time. The history of observations at any $t \geq 0$ is

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

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 $r R > 0$ 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 the default, etc.). Given this assumption, experimentation after the first failure becomes non-profitable, so the players stop experimenting no later than 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 player gets the value of the outside option, and the game is over.

Due to the above assumption, if $O_t \ne \emptyset$, the game is over. Since any player *remaining* 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 $(\Omega^j, \mathcal{F}^j, \{\mathcal{F}_s^j\}_{s \ge 0})$ 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^0 taking values in $[0, 1]$, adapted to the filtration $\{\mathcal{F}_s^j\}_{s \ge 0}$, $j \ne i$, with non-increasing, left-continuous with right limits (LCRL) trajectories; and $q_i^0(0) = 1$.

Note that $q_i^0(t)$ is the probability that player i will not stop at time t or earlier, *conditional* on no observations up to time t. We allow for the case $q_i^0(+\infty) > 0$ which means that player i will not stop ever with the positive probability $q_i^0(+\infty)$ unless a piece of news arrives. Thus, the probability that the player will not stop is $q_i^0(+\infty)(1 - \pi_0)$.

Denote by $\mathcal{H}_t$ the set of histories up to time t, s.t. $O_t = \emptyset$. Following Laraki et al. (2005), and Dutta and Rustichini (1993), for any time $t > 0$, define a proper subgame as the timing game that starts at the end of the history $h_t \in \mathcal{H}_t$.

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

Here $q_i^{t,h}(s)$ is the probability that player i will not exit during $[t, s]$, conditional on being active at t and no observations up to time s, given the history $h \in \mathcal{H}_t$. As before, $q_i^{t,h}(+\infty) > 0$ is allowed with the interpretation that player i who was still

active at time t after the history h_t will never stop with probability $q_i^{t,h}(+\infty)$ unless a piece of news arrives. Thus, the probability that the player who is active at time t will not stop is $q_i^{t,h}(+\infty)(1-\pi_0)$.

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 $h'|_{[0,t]} = h$,

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

The consistency condition above admits interpretation similar to the one in Laraki et al. (2005).[6] The consistency condition asserts that as long as player i is active at t with probability one, later strategies can be calculated by the Bayes rule. With slight abuse of notation, we write q_i^t instead of $q_i^{t,h}$.

3.2 Value functions and equilibrium

Let $G(t)$ denote the instantaneous expected payoff flow of player i if none of the players stopped until time $t > 0$. Let $F(t)$ denote the expected (time t) payoff 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.

For any time $t \ge 0$ consider the subgame (game itself if $t = 0$) after the history $h_t \in \mathcal{H}_t$ such that no news arrived and both players are still active. In all the considerations below, we suppress dependence on parameters of the model and keep dependence on the priors only where necessary.

Given the strategy profile (q_i^t, q_j^t), the value of player i in the subgame game that starts at $t \ge 0$ is

$$
\begin{aligned}
V_i(t; q_i^t, q_j^t) &= \int_t^\infty e^{-rt('-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} G(t') q_i^t(t') q_j^t(t') dt' \\
&+ \int_{\{t' \ge t \,|\, \Delta q_i(t')=0\}} e^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} P(t') q_i^t(t')(-dq_j^t(t')) \\
&+ \int_t^\infty e^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} S q_j(t')(-dq_i^t(t')), \tag{3.2}
\end{aligned}
$$

where the integrals above are the Riemann–Stieltjes integrals as defined in Kolmogorov and Fomin (1975, Section 36.4) . Here the first integral is the expected present value of the flow payoffs $G(t')$ conditioned on both players being active and on no news arrival before time t'. The second integral is 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 $G(t')$ and $P(t')$ are continuous and have finite limits as $t' \to \infty$, hence, $V_i(t; q_i^t, q_j^t)$ is well defined and finite.

---

[6] In the aforementioned paper, strategies of players in a subgame that starts at $t \ge 0$ are probabilities of exit in the interval $[t, s)$.

Note that the second integral in (3.2) takes into account jumps in q_j^t only, and the last integral takes into account jumps in q_i^t only as well as simultaneous jumps in q_i^t and q_j^t.

Definition 3.4 A strategy profile $\hat{q}^0 = (\hat{q}_i^0, \hat{q}_j^0)$ is a Nash equilibrium for the game starting at $t = 0$, if for every $(i, j) \in \{(1, 2), (2, 1)\}$

$$\mathcal{V}_i(0; \hat{q}_j^0) = \sup_{q_i^0} V_i(0; q_i^0, \hat{q}_j^0).$$

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 $t \geq 0$, $\hat{q}^t$ is a Nash equilibrium in the subgame that starts at $t > 0$ (when payoffs are discounted to time t).

3.3 Reformulation

Boyarchenko (2018) shows that if the stopping time game described in the previous section is symmetric (i.e., exchangeable against permutations of the players), then the equilibrium strategies (on the equilibrium path) are of the following type

$$q_i^t(t') = \begin{cases} 1 & \text{if } t \leq t' \leq T_i, \\ 0 & \text{if } t \leq T_i < t'. \end{cases} \tag{3.3}$$

If $T_i = \infty$, then player i never exits. In Sects. 4.3 and 5.3, we also define off the equilibrium path strategies, i.e., strategies in subgames that start at $t > T_i$, provided T_i is finite.

In the present paper, we look for an equilibrium in the class of threshold strategies (3.3).[7] Thus, with slight abuse of notation, from now on, we will consider value functions $V_i(t; T_i, T_j)$ instead of $V_i(t; q_i^t, q_j^t)$. Then, instead of (3.2), we can write

$$
\begin{aligned}
V_i(t; T_i, T_j) = & \int_t^{T_i \wedge T_j} e^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} G(t') dt' \\
& + \mathbb{1}_{T_j < T_i} e^{-r(T_j - t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} F(T_i, T_j) \\
& + \mathbb{1}_{T_i \leq T_j} e^{-r(T_i - t)} \frac{p_2(\pi_0; T_i)}{p_2(\pi_0; t)} S,
\end{aligned}
\tag{3.4}
$$

---

[7] It is straightforward to show that a mixed strategies equilibrium does not exist in either good or bad news model. Indeed, it follows from the Bellman equation that in the region, where $\mathcal{V}_i(t; 1) = S$ (that is the region, where player i is indifferent between stopping and continuation, provided player j is still active), $\Delta q_j^t / \Delta t \geq 0$, but this contradicts Definition 3.2.

where $F(T_i, T_j)$ is the value player i gets at the moment T_j, when player j stops, but player i plans to stay until $T_i \geq T_j$

3.4 Conclusive failures—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 $\Lambda_2(\pi_0; t)$ 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 latter outcome is possible only if the ratio

$$A = \frac{r(R - S)}{C} \tag{3.5}$$

is sufficiently small.

Once one of the players has *quit* experimentation, the other player faces a non-strategic stopping problem, which can be easily solved. Consider a subgame that starts after the history such that no news arrived, and only player j has stopped at time T_j. If player i plans to stop at time $T_i \geq T_j$, her payoff function (at time T_j) is

$$F(T_i, T_j) = \int_{T_j}^{T_i} e^{-r(t'-T_j)} \frac{p_1(\pi_1(\pi_0; T_j); t')}{p_1(\pi_1(\pi_0; T_j); T_j)} \left(rR + \Lambda_1(\pi_1(\pi_0; T_j); t')(S - C)\right) dt'$$
$$+ S e^{-r(T_i - T_j)} \frac{p_1(\pi_1(\pi_0; T_j), T_i)}{p_1(\pi_1(\pi_0; T_j); T_j)}. \tag{3.6}$$

Let A be given by (3.5) and $T \geq t$. Define

$$\Phi(A, t; T) = C \int_t^T e^{-r(t'-t)} \frac{p_1(\pi_1(\pi_0; t); t')}{p_1(\pi_1(\pi_0; t); t)} \left(A - \Lambda_1(\pi_1(\pi_0; t); t')\right) dt'. \tag{3.7}$$

Lemma 3.5 *The payoff function* (3.6) *can be written as*

$$F(T_i, T_j) = S + \Phi(A, T_j; T_i). \tag{3.8}$$

The first term in representation (3.8) is the value of immediate exit; the second term is the value of waiting. See Sect. 8.3 for the proof.

Substituting (3.8) into (3.4), we get

$$V_i(t; T_i, T_j) = \int_t^{T_i \wedge T_j} \mathrm{e}^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} G(t') \mathrm{d}t'$$

$$+ \mathbb{1}_{T_j < T_i} \mathrm{e}^{-r(T_j-t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \Phi(A; T_j; T_i)$$

$$+ \mathrm{e}^{-r(T_i \wedge T_j - t)} \frac{p_2(\pi_0; T_i \wedge T_j)}{p_2(\pi_0; t)} S. \tag{3.9}$$

Assuming that, if the project is bad, the players are equally likely to incur a costly failure, we can write $G(t') = rR + \Lambda_2(\pi_0; t')(S - 0.5C)$.

For $T \geq t$, define

$$\Psi(A, t; T) = C \int_t^T \mathrm{e}^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} \left(A - 0.5\Lambda_2(\pi_0; t') \right) \mathrm{d}t'. \tag{3.10}$$

Lemma 3.6 *The payoff function* (3.9) *can be written as*

$$V_i(t; T_i, T_j) = S + \Psi(A, t; T_i \wedge T_j) + \mathbb{1}_{T_j < T_i} \mathrm{e}^{-r(T_j-t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \Phi(A, T_j; T_i). \tag{3.11}$$

See Sect. 8.4 for the proof.

3.5 Conclusive breakthroughs: good news model

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 other player gets the outside option. We will indicate later how the results may change if the winner takes the major prize, but the other player is also rewarded with the reward being higher than the value of the outside option. Assume that two symmetric players experiment with technologies of unknown quality. Experimentation is costly, and the stream of experimentation costs is rC for each player, independently of the quality of the project. For simplicity, assume that if the project if good, active players stop experimentation after the first success has been observed. Further assume that, if the project is good, the player who is first to succeed gets $R > 0$, and the other player gets $S \in [0, R)$. 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 also gets S. We use the same primitives as described in Sect. 2 and same strategies as in Sect. 3.1.

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 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, for example, due to the deadline specified by designers of the grant competition. To this end, we assume that

$$\sup_{T>0} \left(\int_0^T e^{-rt} p_2(\pi_0; t) \left((R+S)0.5\Lambda_2(\pi_0; t) - rC \right) dt + e^{-rT} p_2(\pi_0; T)S \right) > S.$$

(3.12)

Then it is optimal for two players to experiment until the time $\hat{t}_2(\pi_0)$ when the rate of arrival reaches its maximal value and for some time after that. As the rate of arrival $\Lambda_2(\pi_0; t)$ starts decreasing, it may become optimal for one or more players to quit. We will show that if 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.

Let $\tilde{G}_i(t) = 0.5\Lambda_2(\pi_0; t)(R+S) - rC$ denote the instantaneous expected payoff flow of player i if none of the players stopped until time $t > 0$. Let $\tilde{F}(T_i, T_j)$ denote the value that player i gets if player j stopped at time $T_j < T_i$ and no news arrived before T_j. Consider a subgame that starts after the history such that no observations arrived, and only player j has stopped at time T_j. If player i plans to stop at time $T_i \geq T_j$, her payoff function (at time T_j) is

$$\tilde{F}(T_i, T_j) = \int_{T_j}^{T_i} e^{-r(t'-T_j)} q_f^t(t') \frac{p_1(\pi_1(\pi_0; T_j); t')}{p_1(\pi_1(\pi_0; T_j); T_j)} \left(R\Lambda_1(\pi_1(\pi_0; T_j); t') - C) \right) dt'$$

$$+ Se^{-r(T_i-T_j)} \frac{p_1(\pi_1(\pi_0; T_j), T_i)}{p_1(\pi_1(\pi_0; T_j); T_j)}.$$

(3.13)

Set $\tilde{A} = r(C+S)/(R-S)$ and, for $T \geq t$, define

$$\tilde{\Phi}(\tilde{A}, t; T) = (R-S) \int_t^T e^{-r(t'-t)} \frac{p_1(\pi_1(\pi_0; t); t')}{p_1(\pi_1(\pi_0; t); t)} \left(\Lambda_1(\pi_1(\pi_0; t); t') - \tilde{A} \right) dt';$$

(3.14)

$$\tilde{\Psi}(\tilde{A}, t; T) = (R-S) \int_t^T e^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} \left(0.5\Lambda_2(\pi_0; t') - \tilde{A} \right) dt'.$$

(3.15)

Following the proof of Lemma 3.5 word by word, one can show that

$$\tilde{F}(T_i, T_j) = S + \tilde{\Phi}(\tilde{A}, T_j; T_i).$$

(3.16)

Next, following exactly the same reasoning as in the proof of Lemma 3.6, it is straightforward to show that the value of player i, given by Eq. (3.4), can be equivalently written as

$$\tilde{V}_i(t; T_i, T_j) = S + \tilde{\Psi}(\tilde{A}, t; T_i \wedge T_j) + \mathbb{1}_{T_j < T_i} e^{-r(T_j - t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \tilde{\Phi}(\tilde{A}, T_j; T_i).$$

(3.17)

3.6 Super(sub)modularity of the stopping time games with random observations

We start with the following result. Introduce $\underline{T} = \underline{T}(A)$ given by

$$A = \Lambda_1(1, \underline{T}),$$

(3.18)

and

$$\phi(A; T, t) = p_2(\pi_0; T) \cdot \frac{p_1(\pi_1(\pi_0; T); t)}{p_1(\pi_1(\pi_0; T); T)} (A - \Lambda_1(\pi_1(\pi_0; T); t)).$$

(3.19)

Lemma 3.7 *Function ϕ is increasing in T on $\{t \geq \underline{T} \vee T\}$.*

See Sect. 8.5 for the proof.

Theorem 3.8 *Restrictions of the payoff functions $V_i(t; \cdot, \cdot)$ on $[t \vee \underline{T}, +\infty)^2$ are supermodular.*

Proof We need to check that for all $T_i' \geq T_i$ and $T_j' \geq T_j$,

$$V_i(t; T_i', T_j') - V_i(t; T_i, T_j') \geq V_i(t; T_i', T_j) - V_i(t; T_i, T_j).$$

(3.20)

See, e.g., Amir (1996b) or Vives (1999) for definition of supermodularity.

To prove that (3.20) holds, we notice that for any fixed T_j, the partial derivative $\partial V_i(t; T_i, T_j)/\partial T_i$ is continuous and non-decreasing in T_j. Moreover, $\partial^2 V_i(t; T_i, T_j)/\partial T_i \partial T_j$ exists on $\{T_i \neq T_j\}$. Indeed, if $T_i < T_j$,

$$\frac{\partial V_i(t; T_i, T_j)}{\partial T_i} = C e^{-r(T_i - t)} \frac{p_2(\pi_0; T_i)}{p_2(\pi_0; t)} (A - 0.5\Lambda_2(\pi_0; T_i)),$$

(3.21)

which is independent of T_j, hence $\partial^2 V_i(t; T_i, T_j)/\partial T_i \partial T_j = 0$ if $T_i < T_j$. Consider the case $T_j < T_i$ and rewrite Eq. (3.11) as

$$V_i(t; T_i, T_j) = S + C \int_t^{T_j} e^{-r(t'-t)} \frac{p_2(\pi_0; t')}{p_2(\pi_0; t)} (A - 0.5\Lambda_2(\pi_0; t')) dt'$$
$$+ \frac{C}{p_2(\pi_0; t)} \int_{T_j}^{T_i} e^{-r(t'-t)} \phi(A; T_j, t') dt',$$

where $\phi(A; T_j, t')$ is given by (3.19). Calculate

$$\frac{\partial V_i(t; T_i, T_j)}{\partial T_i} = \frac{C}{p_2(\pi_0; t)} e^{-r(T_i - t)} \phi(A; T_j, T_i).$$

(3.22)

By Lemma 2.2, $\phi(A; T_i, T_i) = p_2(\pi_0; T_i)(A - 0.5\Lambda_2(\pi_0; T_i))$, hence $\partial V_i(t; T_i, T_j)/\partial T_i$ is continuous at $T_i = T_j$. By Lemma 3.7, function $\phi(A; T, t)$ is increasing in T on $\{t \geq \underline{T} \vee T\}$. Hence, $\partial^2 V_i(t; T_i, T_j)/\partial T_i \partial T_j \geq 0$ if $\underline{T} < T_j < T_i$.

The increasing differences condition (3.20) is equivalent to

$$\int_{T_i}^{T_i'} \left(\frac{\partial V_i(t; \tau, T_j')}{\partial \tau} - \frac{\partial V_i(t; \tau, T_j)}{\partial \tau} \right) d\tau \geq 0. \tag{3.23}$$

In the regions $\tau > T_j' \geq T_j > \underline{T}$ and $\underline{T} < \tau < T_j \leq T_j'$, $\partial^2 V_i(t; \tau, T_j)\partial \tau \partial T_j \geq 0$. Applying the mean value theorem, we conclude that

$$\frac{\partial V_i(t; \tau, T_j')}{\partial \tau} - \frac{\partial V_i(t; \tau, T_j)}{\partial \tau} \geq 0 \tag{3.24}$$

if either $\tau \geq T_j' \geq T_j \geq \underline{T}$ or $\underline{T} \leq \tau \leq T_j \leq T_j'$. If $\underline{T} \leq T_j < \tau < T_j'$, we write (3.24) as

$$\frac{\partial V_i(t; \tau, T_j')}{\partial \tau} - \frac{\partial V_i(t; \tau, \tau)}{\partial \tau} + \frac{\partial V_i(t; \tau, \tau)}{\partial \tau} - \frac{\partial V_i(t; \tau, T_j)}{\partial \tau} \geq 0 + 0 = 0.$$

Hence, (3.24) for all $\tau \in [T_i, T_i']$. Integrating (3.24), we obtain (3.23). $\square$

Let $\tilde{T} = \tilde{T}(\tilde{A})$ be given by $\tilde{A} = \Lambda_1(1, \tilde{T})$.

Corollary 3.9 *Restrictions of the payoff functions $\tilde{V}_i(t; \cdot, \cdot)$ on $[t \vee \tilde{T}, +\infty)^2$ are submodular.*

Proof We need to show that the functions $\tilde{V}_i(t; \cdot, \cdot)$ satisfy the decreasing differences condition (i.e., inequality in (3.20) holds with the opposite sign if V_i is replaced by $\tilde{V}_i$). To this end, in the proof of Theorem (3.8), we replace A, V_i, $\underline{T}$ with $\tilde{A}$, $\tilde{V}_i$, $\tilde{T}$, respectively, and change the signs in the flow payoffs, inequalities for the partial derivatives and integrals. $\square$

4 Equilibria in the game with conclusive failures

4.1 Value of the leader

Suppose player j pre-commits to stop no earlier than player i. Then, at any $t \geq 0$, player i is the leader and solves the following problem:

$$\sup_{T_i \geq t} \Psi(A, t; T_i),$$

and the leader's value function is

$$V_L(t) = S + \sup_{T_i \geq t} \Psi(A, t; T_i). \tag{4.1}$$

Let

$$\hat{A}_2 = \hat{A}_2(\pi_0) = \frac{1}{2} \max_{t \geq 0} \Lambda_2(\pi_0; t) = \frac{1}{2} \Lambda_2(\pi_0; \hat{t}_2(\pi_0)). \tag{4.2}$$

If $A \geq \hat{A}_2$, then the integrand in (3.10) is nonnegative (positive for some t'), hence $\Psi(A, t; T_i)$ is strictly increasing in T_i, therefore player i never exits before the first failure happens, and so does player j, hence a unique Nash equilibrium is $T_i^*(t) = T_j^*(t) = \infty$ for all $t \geq 0$.

Lemma 4.1 *Let $A < \hat{A}_2(\pi_0)$, then*

(a) *the equation*

$$A - 0.5\Lambda_2(\pi_0; t) = 0 \tag{4.3}$$

   has two solutions $t_2^*(A, \pi_0) < t_{*2}(A, \pi_0)$;
(b) *$t_2^*(A, \pi_0)$ is the local maximum, and $t_{*2}(A, \pi_0)$ is the local minimum of $\Psi(A, 0; \cdot)$.*

Proof (a) By Lemma 2.3, $\Lambda_2(\pi_0; t)$ is hump-shaped. Therefore, if $A < \hat{A}_2(\pi_0)$ equation (4.3) has exactly two solutions:

$$0 < t_2^*(A, \pi_0) < \hat{t}_2(\pi_0) < t_{*2}(A, \pi_0).$$

(b) If $0 < t_2^*(A, \pi_0) < t_{*2}(A, \pi_0)$ are solutions to (4.3), then it is easy to see that the LHS in (4.3) is positive if $t < t_2^*$ or $t > t_{*2}$; and it is negative if $t_2^* < t < t_{*2}$. Since $\Psi(A, 0; \cdot)$ is increasing (respectively, decreasing) iff the LHS in (4.3) is positive (respectively, negative), (b) follows. $\qquad\qquad\square$

Notice that in the correspondent exponential model with parameter λ, the FOC analogous to (4.3) is $A - \lambda\pi_2(\pi_0; t) = 0$; and the leader's value function is increasing (respectively, decreasing) if $A - \lambda\pi_2(\pi_0; t) > 0$ (respectively, $A - \lambda\pi_2(\pi_0; t) < 0$). If $\lambda\pi_2(\pi_0; 0)) < A$ it is not optimal to start experimentation, because the leader's value function is negative. If $\lambda\pi_2(\pi_0; 0)) > A$, there exists a unique solution to the FOC, but it is the point of minimum of the leader's value function. That is why in the exponential model, players never stop before the first breakdown happens. See Fig. 1a for illustration.

The next result indicates that if A is sufficiently small, the local maximum $t_2^*(A, \pi_0)$ of $\Psi(A, 0, \cdot)$ becomes the global maximum.

Lemma 4.2 *There exists a unique $A_2^* = A_2^*(\pi_0) \in (0, \hat{A}_2(\pi_0))$ s.t.*

$$\Psi(A_2^*, t_2^*; +\infty) = 0, \tag{4.4}$$

and $\Psi(A, t_2^*; +\infty) > 0 (= 0, < 0)$ if $A > A_2^* (= A_2^*, < A_2^*)$.
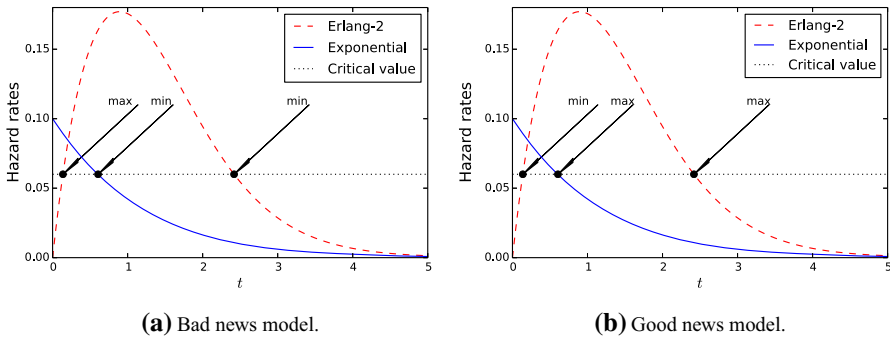
(a) Bad news model.

(b) Good news model.

Fig. 1 Hazard rates: Erlang-2 versus exponential model

Proof Fix π_0, and suppress the dependence on π_0 in the notation $\hat{A}_2$, $t_2^*(A)$, A_2^*. For $A < \hat{A}_2$, consider

$$\Psi(A, t_2^*(A); +\infty) = \frac{Ce^{rt_2^*}}{p_2(\pi_0; t_2^*(A))} \int_{t_2^*(A)}^{\infty} e^{-rt} p_2(\pi_0; t) (A - 0.5\Lambda_2(\pi_0; t)) \, dt. \tag{4.5}$$

The integrand in (4.5) increases in A. Furthermore, while A remains below $\hat{A}_2$, the integrand is negative in a right neighborhood of $t_2^*(A)$, and $t_2^*(A)$ moves to the right as A increases. Hence, the integral in (4.5) is increasing in A. As a function of A, the integral is positive at $\hat{A}_2$; by continuity, it is also positive in a left neighborhood of $\hat{A}_2$. In the limit $A \to 0+$, the integral becomes negative. Hence, there exists $A_2^* \in (0, \hat{A}_2)$ s.t. the integral in (4.5) is negative for any $A < A_2^*$, and positive for any $A \in (A_2^*, \hat{A}_2)$. By monotonicity of $\Psi(A, t_2^*(A); +\infty)$ w.r.t. $t_2^*(A)$, this A_2^* is unique. □

It follows from Lemma 4.2, that for all $A > A_2^*(\pi_0)$, the leader never stops before the first failure, hence the follower also does not stop before the first failure, hence the unique equilibrium is $\hat{q}_i^t(t') = \hat{q}_j^t(t') = 1$ for all $t' > t \geq 0$. This equilibrium is a SPE, because in every subgame that starts after the history $h_t \in \mathcal{H}_t$ s.t. both players are active and no failures were observed, $\arg\max_{T \geq t} \Psi(A, t; T) = +\infty$. In a non-generic case $A = A_2^*(\pi_0)$, the leader is indifferent between stopping at $t_2^*(A_2^*, \pi_0)$ and never stopping before the first failure.

If $A < A_2^*$, then $t_2^*(A, \pi_0)$ is the global maximum of $\Psi(A, 0; \cdot)$, because $\Psi(A, t_2^*; \infty) < 0$. Moreover, since $\Psi(A, t_{*2}; \infty) > 0$ and $\Psi(A, \cdot; \infty)$ increases on $(t_2^*(A, \pi_0), t_{*2}(A, \pi_0))$ there exists a unique solution to $\Psi(A, t; +\infty) = 0$. Denote this solution $\bar{T} = \bar{T}(A, \pi_0)$. Clearly, $\bar{T} \in (t_2^*(A, \pi_0), t_{*2}(A, \pi_0))$, and for $t < \bar{T}$, $\Psi(A, t; +\infty) < 0$, and for $t > \bar{T}$, $\Psi(A, t; +\infty) > 0$.

It follows from the above analysis that the leader's optimal strategy in any subgame that starts at time $t \geq 0$ is

$$T_i^*(t) = \begin{cases} t_2^*(A, \pi_0), & \text{if } 0 \le t \le t_2^*(A, \pi_0), \\ t, & \text{if } t_2^*(A, \pi_0) < t \le \bar{T}, \\ \infty, & \text{if } t > \bar{T}. \end{cases}$$

4.2 Value of the follower

Suppose player i pre-commits to stop no earlier than player j. Then player i is the follower who solves the following problem:

$$\sup_{T_i \ge T_j} \Phi(A, T_j; T_i),$$

and the follower's value function after the history $h_t(\emptyset, \infty, \infty)$ is

$$V_F(t; T_j) = S + \Psi(A, t; T_j) + e^{-r(T_j - t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \sup_{T_i \ge T_j} \Phi(A, T_j; T_i). \quad (4.6)$$

From the previous section, we know that if $A > A_2^*$, both the leader and the follower do not stop before the first failure is observed, therefore from now on, we focus on the case $A < A_2^*$.

Consider $t_2^*(A, \pi_0) \le T_j \le \bar{T}$. (If $T_j > \bar{T}$, then we know from the previous section that the players do not stop before the first observation arrives.) On the strength of inequality (8.5) $\Psi(A, t_2^*(A, \pi_0); +\infty) < 0$ implies $\Phi(A, T_j; +\infty) < 0$ for any $t_2^*(A, \pi_0) \le T_j \le \bar{T}$. Therefore, the best response of the follower is $T_i^*(T_j) = T_j$ for any $t_2^*(A, \pi_0) \le T_j \le \bar{T}$, i.e., the follower stops together with the leader.

It remains to consider the case $0 \le T_j < t_2^*(A, \pi_0)$.

Theorem 4.3 *Let $\Psi(A, t_2^*(A, \pi_0); +\infty) < 0$. Then, in any subgame that starts at $0 \le T_j < t_2^*(A, \pi_0)$ after the history $h_{T_j}(\emptyset, \infty, T_j)$, the follower's best response is $T_i^*(T_j) = t_1^*(A, \pi_1(\pi_0; T_j))$, where $t_1^*(A, \pi_1(\pi_0; T_j)) \in (T_j, t_2^*(A, \pi_0))$ is the smallest solution to*

$$A - \Lambda_1(\pi_1(\pi_0; T_j); t) = 0. \quad (4.7)$$

Moreover $t_1^*(A, \pi_1(\pi_0; T_j))$ is increasing in T_j.

See Sect. 8.6 for the proof. To summarize, the follower's best response function in any subgame that starts after the history $h_{T_j}(\emptyset, \infty, T_j)$ is

$$T_i^*(T_j) = \begin{cases} t_1^*(A, \pi_1(\pi_0; T_j)), & \text{if } T_j < t_2^*(A, \pi_0), \\ T_j & \text{if } T_j \ge t_2^*(A, \pi_0). \end{cases} \quad (4.8)$$

The follower's value after the history $h_t(\emptyset, \infty, \infty)$ is

(a) Best response function of player 1.

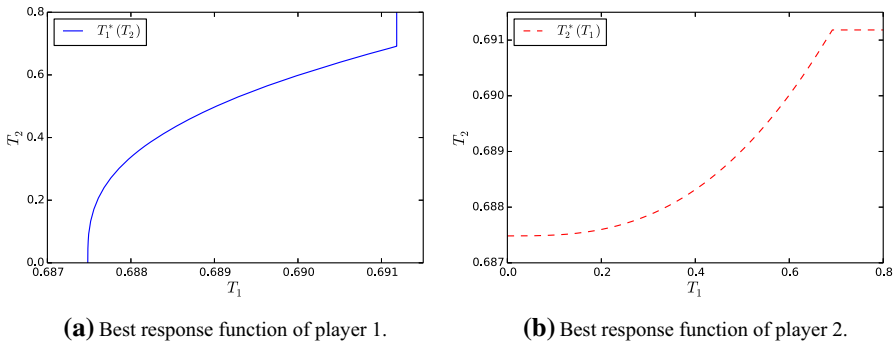**(b)** Best response function of player 2.

Fig. 2 Players' best response functions in the bad news model

$$V_F(t; T_j) = S + \Psi(A, t; T_j) + \mathbb{1}_{T_j < t_2^*(A, \pi_0)} e^{-r(T_j - t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \Phi(A, T_j; t_1^*(A, \pi_1(\pi_0; T_j))).$$

Thus, for any $T_j \geq t_2^*(A, \pi_0)$, $V_F(t; T_j) = S + \Psi(A, t; T_j)$, i.e., the follower's value equals the leader's value. For any $T_j < t_2^*(A, \pi_0)$, the follower's value is greater than the value of the player who exits at T_j, because $\Phi(A, T_j; t_1^*(A, \pi_1(\pi_0; T_j))) > 0$. Hence, the interval $[0, t_2^*(A, \pi_0))$ plays the role of an attrition zone.

It follows from the above analysis that if player j stops at $T_j < t_2^*(A, \pi_0)$, then player i becomes the follower, who stops at $t_1^*(A, \pi_1(\pi_0; T_j))$. If player j stops at $T_j \geq t_2^*(A, \pi_0)$, then stopping at $t_2^*(A, \pi_0)$ is the dominant strategy for player i. Hence, in the game without precommitment, the best response functions of the players are

$$T_i^*(T_j) = \begin{cases} t_1^*(A, \pi_1(\pi_0; T_j)), & \text{if } T_j < t_2^*(A, \pi_0), \\ t_2^*(A, \pi_0) & \text{if } T_j \geq t_2^*(A, \pi_0). \end{cases} \tag{4.9}$$

See Fig. 2 for illustration. Notice that

$$t_1^*(A, \pi_1(\pi_0; t_2^*(A, \pi_0))) = t_2^*(A, \pi_0),$$

because the RHS is the smallest solution to (4.3), the LHS is the smallest solution to (4.7); functions $\Lambda_1(\pi_1(\pi_0, t_2^*(\pi_0, A)); t)$ and $0.5\Lambda_2(\pi_0; t)$ have only one positive point of intersection at $t = t_2^*(\pi_0, A)$. Thus, (i) $T_i^*(T_j) = t_1^*(A, \pi_1(\pi_0; T_j))$ is increasing in the second argument and $T_i^*(T_j) > T_j$ for all $T_j \in [0, t_2^*(A, \pi_0))$ (ii) $T_i^*(T_j) = t_2^*(A, \pi_0)$ for all $T_j \geq t_2^*(A, \pi_0)$. Therefore, the point

$$T_i^*(T_j) = t_1^*(A, \pi_1(\pi_0; T_j)) = t_2^*(A, \pi_0) = T_j$$

is unique. Hence, the unique equilibrium outcome is that the leader and the follower stop together at $T_i = T_j = t_2^*(A, \pi_0)$.

4.3 Equilibrium

Theorem 4.4 *Let $A < A^*(\pi_0)$. Then the stopping time game with the payoff functions defined by* (3.11) *has a unique SPE given by the following pair of consistent strategies:*

$$\hat{q}_i^t(t') = \begin{cases} 1, & \text{if } 0 \le t \le t' \le t_2^*(A, \pi_0), \\ 0, & \text{if } 0 \le t \le t_2^*(A, \pi_0) < t', \\ & \text{or } t_2^*(A, \pi_0) < t \le \min\{t', \bar{T}\}, \\ 1, & \text{if } \bar{T} \le t \le t'. \end{cases} \qquad i \in \{1, 2\} \qquad (4.10)$$

The equilibrium payoffs are

$$V_i(t; \hat{q}_i^t, \hat{q}_j^t) = \begin{cases} S + \Psi(A, t; t_2^*(A, \pi_0)), & \text{if } t \in [0, t_2^*(A, \pi_0)), \\ S, & \text{if } t \in [t_2^*(A, \pi_0), \bar{T}), \\ S + \Psi(A, t; \infty), & \text{if } t \ge \bar{T}. \end{cases} \qquad i \in \{1, 2\}$$

Proof Recall that we look for a NE in the class of strategies of the threshold type specified by (3.3). Therefore, there is one-to-one correspondence between q_i^t and the point $T_i \ge t$ s.t. $q^t(t') = 0$ for all $t' > T_i$. The equilibrium exists and it is unique, because the best response functions $T_1^*(T_2)$ and $T_2^*(T_1)$ intersect, and they intersect only once. See Fig. 4a for illustration. The strategies (4.10) are consistent, hence the proposed strategy profile is SPE.   □

Remark 4.5 Notice that on the equilibrium path, both players stop at t_2^*. Off the equilibrium pass, they may be still active in a subgame that starts at $t > t_2^*$. If $t < \bar{T}$, the players will stop immediately and get the safe payoff S. If $t > \bar{T}$, the players will remain active until the first piece of bad news arrives, therefore, their expected payoffs in a subgame that starts at $t > \bar{T}$ are bigger than the safe payoff S. If $t = \bar{T}$, the players are indifferent between stopping immediately and staying until the first failure happens.

4.4 Efficiency of the equilibrium outcome

First of all, notice that the equilibrium outcome is the same as a cooperative solution, where the two players agree on how long they will stay in the game together, because the cooperative solution maximizes $2(S + \Psi(A, 0; T))$, and the maximum is achieved either at $T = +\infty$ (if $A > A_2^*$) or at $T = t_2^* = t_2^*(A, \pi_0)$ (if $A < A_2^*$). In a non-generic case $A = A_2^*$, the players will be indifferent between stopping at t_2^* and never unless the first failure happens.

Now consider the social planner, who chooses the leader's stopping time T so as to maximize the sum of the values of both players, i.e., $2S + 2\Psi(A, 0; T) + e^{-rT} p_2(\pi_0; T)\Phi(A, T; T_f(T))$, where $T_f(T)$ is the follower's best response to T. Equivalently, the social planner chooses T to maximize the payoff function

$$SP(A; T) = 2S + 2C \int_0^T e^{-rt} p_2(\pi_0; t) (A - 0.5\Lambda_2(\pi_0; t)) \, dt$$

$$+ C \int_T^{T_f(T)} e^{-rt} \phi(A; T, t) dt, \qquad (4.11)$$

where $\phi(A; T, t)$ is defined by (3.19). Recall that if $T \geq t_2^*(A, \pi_0)$, then $T_f(T) = T$, therefore the second integral in (4.11) disappears, and

$$\frac{\partial SP(A; T)}{\partial T} = 2\mathrm{e}^{-rT} p_2(\pi_0; T) (A - 0.5\Lambda_2(\pi_0; T)), \text{ if } T > t_2^*.$$

The above partial derivative is negative if $T \in (t_2^*, t_{*2})$ and positive if $T > t_{*2}$, therefore if a maximum of the social planner's function is attained on $(t_2^*, \infty]$, the maximum is at $T = \infty$. In Sect. 4.1, we showed that the maximum of $\Psi(A; 0, T)$ is attained at $T = \infty$ if $A > A_2^*$. Hence the equilibrium outcome where both players stay until the first piece of bad news arrives is efficient if $A > A_2^*$. The non-generic case $A = A_2^*$ produces efficient outcome by the same argument.

Let $A < A_2^*$. Then the maximum of $\Psi(A; 0, T)$ is attained at $T = t_2^*$, and $T_f(t_2^*) = t_2^*$. Hence $SP(A; t_2^*) > SP(A; \infty)$. To find the maxima of $SP(A; \cdot)$ on $[0, t_2^*]$, calculate the derivative on $(0, t_2^*)$

$$\frac{\partial SP(A; T)}{\partial T} = 2\mathrm{e}^{-rT} p_2(\pi_0; T) (A - 0.5\Lambda_2(\pi_0; T)) + \int_T^{T_f(T)} \mathrm{e}^{-rt} \frac{\partial \phi(A; T, t)}{\partial T} \mathrm{d}t$$

$$- \mathrm{e}^{-rT} \phi(A; T, T) + \mathrm{e}^{-rT_f(T)} \phi(A; T, T_f(T)) \frac{\partial T_f(T)}{\partial T}. \quad (4.12)$$

Recall that if $T \leq t_2^*(A, \pi_0)$, then $A - \Lambda_1(\pi_1(\pi_0; T); T_f) = 0$; in this case [see (3.19)], the last term in (4.12) is zero. Using the property $0.5\Lambda_2(\pi_0; T) = \Lambda_1(\pi_1(\pi_0; T); T)$ and the definition of $\phi(A; T, t)$ (3.19), we simplify (4.12):

$$\frac{\partial SP(A; T)}{\partial T} = \mathrm{e}^{-rT} p_2(\pi_0; T) (A - 0.5\Lambda_2(\pi_0; T)) + \int_T^{T_f(T)} \mathrm{e}^{-rt} \frac{\partial \phi(A; T, t)}{\partial T} \mathrm{d}t. \quad (4.13)$$

The first term is positive if $T \in (0, t_2^*)$. By Lemma 3.7, $\partial \phi(A; T, t)/\partial T \geq 0$ if $T \geq \underline{T}$. Since $t_2^* > \underline{T}$, $\partial SP(A; T)/\partial T > 0$ if $T \in [\underline{T}, t_2^*)$. Evidently, $\partial SP(A; t_2^*-)/\partial T = 0$, hence $T = t_2^*$ is a local maximum of $SP(A; \cdot)$ on $[\underline{T}, \infty)$. Recall that if $T < \underline{T}$, then experimentation is optimal even if the project is bad for sure, therefore the social planner will not find it optimal to stop one of the players at $T \in [0, \underline{T})$. Hence, $T = t_2^*$ is the social planner's solution, i.e., the equilibrium outcome is efficient.

5 Equilibria in the game with conclusive successes

5.1 Value of the leader

Suppose player j pre-commits to stop no earlier than player i. Then, at any $t \geq 0$, player i is the leader who solves the following problem:

$$\sup_{T_i \geq t} \tilde{\Psi}(\tilde{A}, t; T_i),$$

and the leader's value function is

$$\tilde{V}_L(t) = S + \sup_{T_i \geq t} \tilde{\Psi}(\tilde{A}, t; T_i). \tag{5.1}$$

Lemma 5.1 *Set $t = 0$. Condition* (3.12) *is equivalent to*

$$\sup_T \tilde{\Psi}(\tilde{A}, 0; T) > 0. \tag{5.2}$$

See Sect. 8.7 for the proof. If $\tilde{A} \geq \hat{A}_2(\pi_0) = 0.5 \max_t \Lambda_2(\pi_0, t)$, then condition (5.2) cannot hold, because the integrand is negative almost everywhere. Hence, if (5.2) holds, $\tilde{A} < \hat{A}_2(\pi_0)$ and we can repeat the proof of Lemma 4.1 word by word to prove the following result.

Lemma 5.2 *Let $\tilde{A} < \hat{A}_2(\pi_0)$, then*

(a) *the equation*

$$0.5\Lambda_2(\pi_0; t) - \tilde{A} = 0 \tag{5.3}$$

has two solutions $t_2^*(\tilde{A}, \pi_0) < t_{*2}(\tilde{A}, \pi_0)$;

(b) *$t_2^*(\tilde{A}, \pi_0)$ is the local minimum, and $t_{*2}(\tilde{A}, \pi_0)$ is the local maximum of $\tilde{\Psi}(\tilde{A}, 0; \cdot)$.*

We can also state that $t_{*2}(\tilde{A}, \pi_0)$ is the global maximum of $\tilde{\Psi}(\tilde{A}, 0; \cdot)$ due to assumption (5.2). The leader's optimal strategy in any subgame that starts at time $t \geq 0$ is

$$T_i^*(t) = \begin{cases} t_{*2}(A, \pi_0), & \text{if } 0 \leq t \leq t_{*2}(A, \pi_0), \\ t, & \text{if } t_{*2}(A, \pi_0) < t. \end{cases}$$

Notice that in any subgame that starts at $t > t_{*2}(\tilde{A}, \pi_0)$, the optimal strategy is to stop immediately because $\tilde{\Psi}(\tilde{A}, t; T)$ is decreasing in the last argument on $\{T > t_{*2}(\tilde{A}, \pi_0)\}$.

Notice that in the correspondent exponential model with parameter λ, the FOC analogous to (5.3) is $\tilde{A} - \lambda\pi_2(\pi_0; t) = 0$; and the leader's value function is increasing (respectively, decreasing) if $\lambda\pi_2(\pi_0; t) - \tilde{A} > 0$ (respectively, $\lambda\pi_2(\pi_0; t) - \tilde{A} < 0$). If $\lambda\pi_2(\pi_0; 0)) < \tilde{A}$ it is not optimal to start experimentation, because the leader's value function is negative. If $\lambda\pi_2(\pi_0; 0)) > \tilde{A}$, there exists a unique solution to the FOC, and this is the point of maximum of the leader's value function. See Fig. 1b for illustration.

5.2 Value of the follower

Suppose player i pre-commits to stop no earlier than player j. Then player i is the follower who solves the following problem:

$$\sup_{T_i \geq T_j} \tilde{\Phi}(\tilde{A}, T_j; T_i),$$

and the follower's value function after the history $h_t(\emptyset, \infty, \infty)$ is

$$\tilde{V}_F(t; T_j) = S + \tilde{\Psi}(\tilde{A}, t; T_j) + e^{-r(T_j - t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \sup_{T_i \geq T_j} \tilde{\Phi}(\tilde{A}, T_j; T_i). \quad (5.4)$$

Consider $T_j \geq t_{*2}(\tilde{A}, \pi_0)$. Then $0.5\Lambda_2(\pi_0, T_j) \leq \tilde{A}$, and by Lemma 2.2,

$$\Lambda_1(\pi_1(\pi_0; T_j); T_j) = 0.5\Lambda_2(\pi_0; T_j) \leq \tilde{A}.$$

Furthermore, since $t_{*2} = t_{*2}(\tilde{A}, \pi_0)$ is the largest solution to (5.3), $t_{*2}(\tilde{A}, \pi_0) > \hat{t}_2(\pi_0)$, hence $T_j > \hat{t}_2(\pi_0)$. It is possible to show that for all $T_j \geq t_{*2}(\tilde{A}, \pi_0)$, $T_j > \hat{t}_1(\pi_1(\pi_0, T_j))$. Therefore, for all $t > T_j \geq t_{*2}$, $\Lambda_1(\pi_1(\pi_0; T_j); t) < \tilde{A}$, i.e., experimentation is not optimal for the follower if the leader stops at any $T_j \geq t_{2*}(\tilde{A}, \pi_0)$. We conclude that the best response of the follower is $T_i^*(T_j) = T_j$ for any $t' \geq T_j \geq t_{*2}(\tilde{A}, \pi_0)$. It remains to consider the case $0 \leq T_j < t_{*2}(\tilde{A}, \pi_0)$.

Theorem 5.3 *Let condition (5.2) hold. Then, in any subgame that starts at $0 \leq T_j < t_{*2}(\tilde{A}, \pi_0)$ after the history $h_{T_j}(\emptyset, \infty, T_j)$, the follower's best response is $T_i^*(T_j) = t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$, where $t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$ is the largest solution to*

$$\Lambda_1(\pi_1(\pi_0; T_j); t) - \tilde{A} = 0. \quad (5.5)$$

Moreover $t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$ is decreasing in T_j.

See Sect. 8.8 for the proof. To summarize, the follower's best response function in any subgame that starts after the history $h_{T_j}(\emptyset, \infty, T_j)$ is

$$T_i^*(T_j) = \begin{cases} t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j)), & \text{if } T_j < t_{*2}(\tilde{A}, \pi_0), \\ T_j & \text{if } T_j \geq t_{*2}(\tilde{A}, \pi_0). \end{cases} \quad (5.6)$$

The follower's value after the history $h_t(\emptyset, \infty, \infty)$ is

$$\tilde{V}_F(t; T_j) = S + \tilde{\Psi}(\tilde{A}, t; T_j)$$
$$+ \mathbb{1}_{T_j < t_{*2}(\tilde{A}, \pi_0)} e^{-r(T_j - t)} \frac{p_2(\pi_0; T_j)}{p_2(\pi_0; t)} \tilde{\Phi}(\tilde{A}, T_j; t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))).$$

Thus, for any $T_j \geq t_{*2}(\tilde{A}, \pi_0)$, $\tilde{V}_F(t; T_j) = S + \tilde{\Psi}(\tilde{A}, t; T_j)$, i.e., the follower's value equals the leader's value. For any $T_j < t_{*2}(\tilde{A}, \pi_0)$, the follower's value is greater than the value of the player who exits at T_j, because $\tilde{\Phi}(\tilde{A}, T_j; t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))) > 0$. Hence, the interval $[0, t_{*2}(\tilde{A}, \pi_0))$ plays the role of an attrition zone.

It follows form the above analysis that if player j stops at $T_j < t_{*2}(\tilde{A}, \pi_0)$, then player i becomes the follower, who stops at $t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$. If player j stops at $T_j \geq t_{*2}(\tilde{A}, \pi_0)$, then stopping at $t_{*2}(\tilde{A}, \pi_0)$ is the dominant strategy for player i.

(a) Best response function of player 1.   **(b)** Best response function of player 2.
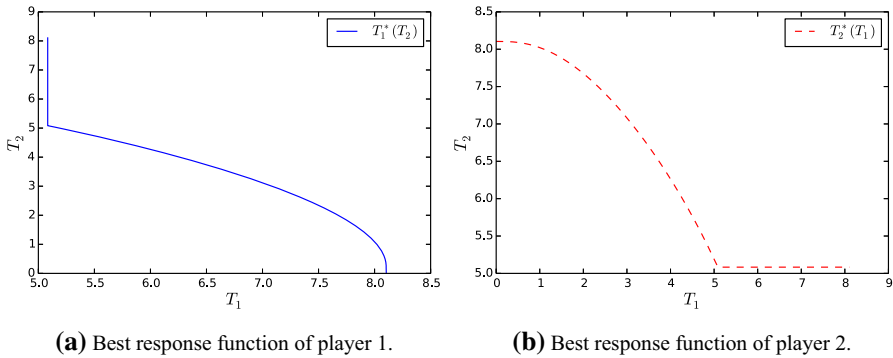
Fig. 3 Players' best response functions in the good news model

Hence, in the game without precommitment, the best response functions of the players are

$$T_i^*(T_j) = \begin{cases} t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j)), & \text{if } T_j < t_{*2}(\tilde{A}, \pi_0), \\ t_{*2}(\tilde{A}, \pi_0) & \text{if } T_j \geq t_{*2}(\tilde{A}, \pi_0). \end{cases} \tag{5.7}$$

See Fig. 3 for illustration. Notice that

$$t_{*1}(\tilde{A}, \pi_1(\pi_0; t_{*2}(\tilde{A}, \pi_0))) = t_{*2}(\tilde{A}, \pi_0),$$

because the RHS is the largest solution to (5.3), the LHS is the largest solution to (5.5); functions $\Lambda_1(\pi_1(\pi_0, t_{*2}(\pi_0, \tilde{A})); t)$ and $0.5\Lambda_2(\pi_0; t)$ have only one positive point of intersection at $t = t_{*2}(\pi_0, \tilde{A})$. Thus, (i) $T_i^*(T_j) = t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$ is decreasing in the second argument and $T_i^*(T_j) > T_j$ for all $T_j \in [0, t_{*2}(\tilde{A}, \pi_0))$ (ii) $T_i^*(T_j) = t_{*2}(\tilde{A}, \pi_0)$ for all $T_j \geq t_{*2}(\tilde{A}, \pi_0)$. Therefore, the point

$$T_i^*(T_j) = t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j)) = t_{*2}(\tilde{A}, \pi_0) = T_j$$

is unique. Hence, the unique equilibrium outcome is that the leader and the follower stop together at $T_i = T_j = t_{*2}(\tilde{A}, \pi_0)$.

5.3 Equilibrium

Theorem 5.4 *Let condition* (5.2) *hold. Then the stopping time game with the value functions defined by* (3.17) *has a unique SPE given by the following pair of consistent strategies:*

$$\hat{q}_i^t(t') = \begin{cases} 1, & \text{if } 0 \leq t \leq t' \leq t_{*2}(\tilde{A}, \pi_0)), \\ 0, & \text{if } 0 \leq t \leq t_{*2}(\tilde{A}, \pi_0) < t', \quad i \in \{1, 2\} \\ & \text{or } t_{*2}(\tilde{A}, \pi_0) < t \leq t'. \end{cases} \tag{5.8}$$
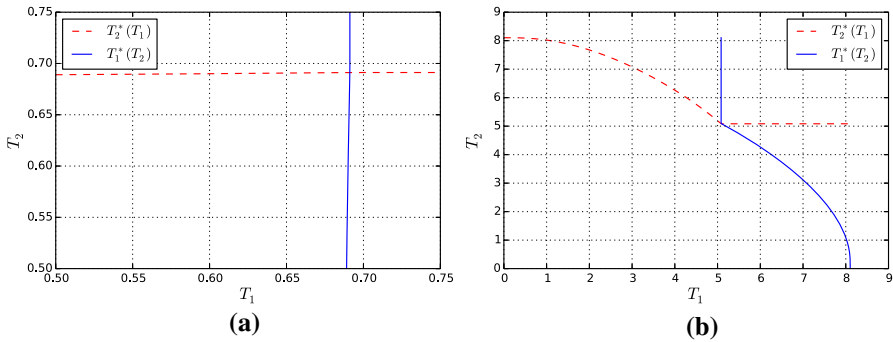
Fig. 4 **a** Equilibrium in the bad news model, **b** equilibrium in the good news model

The equilibrium payoffs are

$$V_i(t; \hat{q}_i^t, \hat{q}_j^t) = \begin{cases} S + \Psi(\tilde{A}, t; t_{*2}(\tilde{A}, \pi_0)), & \text{if } t \in [0, t_{*2}(\tilde{A}, \pi_0)), \\ S, & \text{if } t \geq t_{*2}(\tilde{A}, \pi_0). \end{cases} \quad i \in \{1, 2\}$$

Proof Recall that we look for a NE in the class of strategies of the threshold type specified by (3.3). Therefore, there is one-to-one correspondence between q_i^t and the point $T_i \geq t$ s.t. $q^t(t') = 0$ for all $t' > T_i$. The equilibrium exists and it is unique, because the best response functions $T_1^*(T_2)$ and $T_2^*(T_1)$ intersect, and they intersect only once. See Fig. 4b for illustration. The strategies (5.8) are consistent, hence the proposed strategy profile is SPE. □

The analysis presented in this section indicates that the critical value of the hazard rates at exit is independent of the number of players. In order that the critical values of hazard rates depended on n, one can depart from the winner takes all prize scheme. For example, suppose, that the player, who gets the first breakthrough gets the prize $R > S$, and the second player gets the payoff $S < P < R$. Then each player's expected payoff in case of arrival of good news is $(R + P)/2$. Let, as before, $A = r(C + S)/(R - S)$, and let $\kappa = 1 + (P - S)/(R - S)$. Then the exit threshold for two players is the largest solution $t_{*2} = t_{*2}(\tilde{A}, \pi_0, \kappa)$ to

$$0.5\Lambda_2(\pi_0; t) - \frac{\tilde{A}}{\kappa} = 0, \tag{5.9}$$

provided

$$\int_0^{t_{*2}} p_2(\pi_0; t) \left(0.5\Lambda_2(\pi_0; t) - \frac{\tilde{A}}{\kappa} \right) dt > 0. \tag{5.10}$$

The last inequality is similar to (3.12) (see the proof of Lemma 5.1 in Sect. 8.7 to understand why). If inequality (5.10) is satisfied, then players exit simultaneously at t_{*2}, because $t_{*1}(t_{*2}) < t_{*2}$, where $t_{*1}(t_{*2})$ is the largest solution to

$$\Lambda_1(\pi_1(\pi_0; t_{*2}); t) - \tilde{A} = 0.$$

We see from (5.9) that the critical value of the hazard rates at exit depends on the number of players. Since $\kappa > 1$, the largest solution to (5.9) is larger than the largest solution to (5.3), so the players will experiment longer.

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 holds 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. (2017).

5.4 Inefficiency of the equilibrium outcome

Let condition (5.2) hold. If the two players decide to cooperate and stay in the game together, then they will maximize $2(S + \tilde{\Psi}(\tilde{A}, 0; T))$, hence the will stop together at $T = t_{*2}$ unless the first success happens earlier. Thus, the equilibrium outcome is the same as a cooperative solution.

Now consider the social planner (or sponsor), who pays for the costs of experimentation of each of the active players and who is the claimant to the payoff R in case of a success. The social planner chooses the leader's stopping time T so as to maximize sum of the values of both players, i.e., $2S + 2\tilde{\Psi}(\tilde{A}, 0; T) + e^{-rT} p_2(\pi_0; T)\tilde{\Phi}(\tilde{A}, T; T_f(T))$, where $T_f(T)$ is the follower's best response to T. Equivalently, the social planner chooses T to maximize the payoff function

$$SP(\tilde{A}; T) = 2S + 2(R - S) \int_0^T e^{-rt} p_2(\pi_0; t) \left(0.5\Lambda_2(\pi_0; t) - \tilde{A}\right) dt$$
$$- (R - S) \int_T^{T_f(T)} e^{-rt} \phi(\tilde{A}; T, t) dt, \tag{5.11}$$

where $\phi(\tilde{A}; T, t)$ is defined by (3.19). Recall that if $T \geq t_{*2}(\tilde{A}, \pi_0)$, then by (5.6), $T_f(T) = T$, hence the last integral in (5.11) is zero, and

$$\frac{\partial SP(\tilde{A}; T)}{\partial T} = 2e^{-rT} p_2(\pi_0; T) \left(0.5\Lambda_2(\pi_0; T) - \tilde{A}\right), \text{ if } T > t_{*2}.$$

The above partial derivative is negative if $T > t_{*2}$, therefore, the social planner's function decreases on $(t_{*2}, +\infty)$. To prove that the maxima of $SP(\tilde{A}; \cdot)$ are on the open interval $[0, t_{*2})$, calculate the derivative using the same reasoning as in Sect. 4.4

$$\frac{\partial SP(\tilde{A}; T)}{\partial T} = e^{-rT} p_2(\pi_0; T) \left(0.5\Lambda_2(\pi_0; T) - \tilde{A}\right) - \int_T^{T_f(T)} e^{-rt} \frac{\partial \phi(\tilde{A}; T, t)}{\partial T} dt. \tag{5.12}$$

Tedious straightforward calculations show that

$$\frac{\partial^2 SP}{\partial T^2}(\tilde{A}; t_{*2}-) = \frac{e^{-rt_{*2}} p_2(\pi_0; t_{*2}) 0.5 \partial_t \Lambda_2(\pi_0; t_{*2}) \pi_2(\pi_0; t_{*2}) \partial_t \Lambda_1(1; t_{*2})}{\partial_t \Lambda_1(\pi_1(\pi_0; t_{*2}); t_{*2})}.$$

(5.13)

Since $t_{*2} = T_f(t_{*2})$ is the largest solution to (5.5), t_{*2} is greater than the global maximum of $\Lambda_1(\pi_1(\pi_1; t_{*2}); t)$. Hence by Lemma 2.3, $\partial_t \Lambda_1(\pi_1(\pi_0; t_{*2}); t_{*2}) < 0$. The numerator in (5.13) is also negative, because $\partial_t \Lambda_1(1, t_{*2}) > 0$ and $\partial_t \Lambda_2(\pi_0; t_{*2}) < 0$ since t_{*2} is the largest solution of (5.3). Hence $\partial^2_{TT} SP(\tilde{A}; t_{*2}-) > 0$, therefore $t_{*2}(\tilde{A}, \pi_0)$ is not a point of maximum of $SP(\tilde{A}; T)$.

Thus the equilibrium outcome is not efficient. It follows from the above considerations that the global maximum of the social planner's payoff function is attained either at $T = 0$ or at some $T^* \in (0, t_{*2})$. The former outcome can happen if parameter $\tilde{A} = r(C + S)/(R - S)$ is sufficiently large, i.e., if the opportunity cost of experimentation is relatively high compared to the net benefit from experimentation. In this case, the sponsor will find it optimal to finance a single participant of the research team. If $\tilde{A}$ is sufficiently small, then it is optimal to finance two competitors for a while, stop financing one of the players earlier than in equilibrium and the other one later than in equilibrium.

6 Extensions and generalization

6.1 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. Results concerning super(sub)modularity of the stopping games were independent of the latter assumption. In this section, we generalize our model further and provide a classification of one-humped bandits. As in Sect. 2, the primitive is the rate of arrival of news when the risky project can generate news for sure. Let $n \geq 1$ players experiment with a project of unknown quality, and $\pi_0 \in (0, 1)$ is the common prior which is the probability of the event that the project generates news.

Lemma 6.1 *Let the following conditions hold:*

(i) $\Lambda_1(1; t)$ *is twice differentiable on* $(0, +\infty)$ *and increasing;*

(ii) $$2(\partial_t \Lambda_1(1; t))^2 > \partial^2_t \Lambda_1(1; t) \cdot \Lambda_1(1; t), \ \forall t > 0;$$

(6.1)

(iii) $$\lim_{t \downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1; t)} \right) > n(1 - \pi_0);$$

(6.2)

Then

(a) *as a function of t,* $\Lambda_n(\pi_0; t) = n\Lambda_1(1; t)\pi_n(\pi_0; t)$ *has the global maximum* $\hat{t}_n = \hat{t}_n(\pi_0)$;
(b) *If* (6.1) *holds but* (6.2) *fails, then* $\Lambda_n(\pi_0; t)$ *is a decreasing function on* $\mathbb{R}_+$.

The proof is the same as the proof of Lemma 2.3. However, since we no longer assume that $\Lambda_1(1; 0) = 0$, additional condition (6.2) is needed to show that the derivative $\partial_t \Lambda_n(\pi_0; t)$ changes sign.

Notice that if, for some $1 < \hat{n} \le n$,

$$1 - \pi_0 < \lim_{t \downarrow 0} \partial_t \left(\frac{-1}{\Lambda_1(1; t)} \right) < \hat{n},$$

then the qualitative behavior of the arrival rate $\Lambda_n(\pi_0; t)$ depends on n. Namely, for $n \ge \hat{n}$, the arrival rate $\Lambda_n(\pi_0; t)$ is a decreasing function, and the model is qualitatively the same as the exponential bandits model.

6.2 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 (6.2) 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 $\Lambda_1(1; 0) > 0$, (6.1) 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 (6.1) 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 6.2 Let (6.1) hold.

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

6.3 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: $\exists m \geq 1$ 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) $n \geq m$ players will not start experimenting unless $n - m$ of them exit instantly.

6.4 Multi-humped bandits

We call the model a multi-humped model, if $\Lambda_n(\pi_0; t)$ has more than one point of local maximum. Examples include

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

6.5 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 $\Lambda_k(\pi_0; t)$, which the k players that are in the game face, is defined as follows. For $k = 1, 2, \ldots, n$ and $t \in [t_k, t_{k+1})$,

$$\Lambda_k(\pi_0; t) = \sum_{j=1}^{k} \Lambda_j(\pi_0; t - t_j).$$

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.

7 Conclusion

This paper contributes to the literature on both supermodular games and stopping games with random observations. We demonstrated that games of strategic experimentation with irreversible stopping decisions can be either supermodular or submodular depending on the type of news players may receive.

We suggested a new model for strategic experimentation, where good or bad news arrive at random times which are modeled as jump times of a time-inhomogeneous Poisson process. These models are almost as tractable as exponential bandit models and can incorporate such realistic features as dependence of the expected rate of news arrival on the time elapsed since the start of the experiment. We characterized SPE both in the model with conclusive failures and with conclusive breakthroughs. In particular, we showed that in the bad news model, for a generic set of parameters, there exists a SPE, where the players stop simultaneously before the first piece of bad news arrives.

In the future, we plan to study systematically instances of super(sub)modularity in stopping time games. We also plan to investigate the possibility of generalization of our approach to the case of more than two players. Finally, 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 hazard rates are quite promising and can be used not only as a framework in learning and experimentation models, but in other areas of economics and finance. Uncertainty about rates and times of arrival of observations 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 more realistic than oversimplified reduced form models of defaultable debt with the exponential hazard rate, and the former models are still tractable.

8 Appendix

8.1 Proof of Lemma 2.1

(i) For $n \in \{1, 2\}$, rewrite (2.5) as $\pi_n(\pi_0; t) = 1 - (1 - \pi_0)/p_n(\pi_0; t)$, and notice that $p_n(\pi_0; t) = 1 - \pi_0 + \pi_0 p_1(1, t)^n$ is decreasing in n because $p_1(1; t) < 1$ for all $t > 0$.

(ii) Fix any $t' > 0$. Let $\hat{\pi}$ be a solution to $\pi_n(\pi_0, t') = \pi_{n-1}(\hat{\pi}, t')$, equivalently

$$\frac{\pi_0 p_1(1; t')^n}{1 - \pi_0 + \pi_0 p_1(1; t')^n} = \frac{\hat{\pi} p_1(1; t')^{n-1}}{1 - \hat{\pi} + \hat{\pi} p_1(1; t')^{n-1}}.$$

Hence $\hat{\pi}(1 - \pi_0) p_1(1; t') = (1 - \hat{\pi})\pi_0$, and

$$\hat{\pi} = \frac{\pi_0 p_1(1; t')}{1 - \pi_0 + \pi_0 p_1(1; t')} = \pi_1(\pi_0; t').$$

This proves (2.7).

(iii) Consider the difference

$$D(t) := \pi_2(\pi_0, t) - \pi_1(\pi_1(\pi_0; t'), t).$$

Since $\pi_2(\pi_0, 0) = \pi_0$, and $\pi_1(\pi_1(\pi_0; t'), 0) = \pi_1(\pi_0; t') < \pi_0$, $D(0) > 0$. It follows from (ii) that $t = t'$ is the only zero of $D(t)$. Hence, to prove (2.8), it suffices to show that $D'(t') < 0$. To this end, calculate (using (2.6) and results from (ii))

$$\begin{aligned}
D'(t') &= \partial_t \pi_2(\pi_0, t') - \partial_t \pi_1(\pi_1(\pi_0; t'), t') \\
&= -2\Lambda_1(1; t')\pi_2(\pi_0; t')(1 - \pi_2(\pi_0; t')) \\
&\quad + \Lambda_1(1; t')\pi_1(\pi_1(\pi_0; t'), t')(1 - \pi_1(\pi_1(\pi_0; t'), t')) \\
&= -\Lambda_1(1; t')\pi_2(\pi_0; t')(1 - \pi_2(\pi_0; t')) < 0.
\end{aligned}$$

8.2 Proof of Lemma 2.2

(a) Intersection at $t = 0$ follows from (2.9) and property (i) $\Lambda_1(1; 0) = 0$. By Lemma 2.1 and Eq. (2.9), the only positive point of intersection is t'.
(b) Equation (2.11) follows from (2.9) and the equality in (2.6).
(c) Using (2.11), we derive

$$\frac{1}{n}\partial_t \Lambda_n(\pi_0; 0+) = \pi_0 \partial_t \Lambda_1(1; 0+) > \pi_1(\pi_0; t')\partial_t \Lambda_1(1; 0+)$$

$$= \frac{1}{n-1}\partial_t \Lambda_{n-1}(\pi_1(\pi_0; t'), 0).$$

By continuity, in a right neighborhood of zero,

$$\frac{1}{n}\partial_t \Lambda_n(\pi_0; t) > \frac{1}{n-1}\partial_t \Lambda_{n-1}(\pi_1(\pi_0; t'), t),$$

$$\frac{1}{n}\Lambda_n(\pi_0; t) > \frac{1}{n-1}\Lambda_{n-1}(\pi_1(\pi_0; t'), t).$$

Since there is only one positive intersection of the hazard rates, (2.12) follows.

8.3 Proof of Lemma 3.5

Notice that

$$\int_t^T d\left[e^{-r(t'-t)}\frac{p_1(\pi_1(\pi_0; t); t')}{p_1(\pi_1(\pi_0; t); t)}\right] = e^{-r(T-t)}\frac{p_1(\pi_1(\pi_0; t); T)}{p_1(\pi_1(\pi_0; t); t)} - 1.$$

At the same time,

$$\int_t^T d\left[e^{-r(t'-t)}\frac{p_1(\pi_1(\pi_0; t); t')}{p_1(\pi_1(\pi_0; t); t)}\right]$$

$$= \int_t^T e^{-r(t'-t)}\frac{p_1(\pi_1(\pi_0; t); t')}{p_1(\pi_1(\pi_0; t); t)}\left(-r - \Lambda_1(\pi_1(\pi_0; t); t')\right)dt'.$$

Hence

$$Se^{-r(T_i-T_j)}\frac{p_1(\pi_1(\pi_0; T_j), T_i)}{p_1(\pi_1(\pi_0; T_j); T_j)}$$

$$= S + S\int_t^T e^{-r(t'-t)}\frac{p_1(\pi_1(\pi_0; t); t')}{p_1(\pi_1(\pi_0; t); t)}\left(-r - \Lambda_1(\pi_1(\pi_0; t); t')\right)dt'.$$

Substituting the last expression into (3.6), we arrive at (3.8) .

8.4 Proof of Lemma 3.6

Notice that

$$\int_t^T d\left(e^{-r(t'-t)}\frac{p_2(\pi_0; t')}{p_2(\pi_0; t)}\right) = e^{-r(T-t)}\frac{p_2(\pi_0; T)}{p_2(\pi_0; t)} - 1.$$

At the same time,

$$\int_t^T d\left(e^{-r(t'-t)}\frac{p_2(\pi_0; t')}{p_2(\pi_0; t)}\right) = \int_t^T e^{-r(t'-t)}\frac{p_2(\pi_0; t')}{p_2(\pi_0; t)}\left(-r - \Lambda_2(\pi_0; t')\right) dt'.$$

Hence,

$$e^{-r(T_i \wedge T_j - t)}\frac{p_2(\pi_0; T_i \wedge T_j)}{p_2(\pi_0; t)} S = S + S \int_t^{T_i \wedge T_j} e^{-r(t'-t)}\frac{p_2(\pi_0; t')}{p_2(\pi_0; t)}\left(-r - \Lambda_2(\pi_0; t')\right) dt'.$$

Substituting the last expression into (3.9), we obtain (3.11).

8.5 Proof of Lemma 3.7

Proof Use (2.4) and (2.5) to write

$$\begin{aligned}
p_1(\pi_1(\pi_0; T); t) &= 1 - \pi_1(\pi_0; T) + \pi_1(\pi_0; T)p_1(1; t) \\
&= 1 - \frac{\pi_0 p_1(1; T)}{p_1(\pi_0; T)} + \frac{\pi_0 p_1(1; T)p_1(1; t)}{p_1(\pi_0; T)} \\
&= \frac{1 - \pi_0 + \pi_0 p_1(1; T)p_1(1; t)}{p_1(\pi_0; T)}; \\
p_1(\pi_1(\pi_0; T); T) &= \frac{1 - \pi_0 + \pi_0 p_1(1; T)p_1(1; T)}{p_1(\pi_0; T)} = \frac{p_2(\pi_0; T)}{p_1(\pi_0; T)}.
\end{aligned}$$

Hence

$$\phi(A; T, t) = (1 - \pi_0 + \pi_0 p_1(1; T)p_1(1; t))(A - \Lambda_1(\pi_1(\pi_0; T); t)). \quad (8.1)$$

Use (2.5) and (2.9) to write

$$\begin{aligned}
\Lambda_1(\pi_1(\pi_0; T); t) &= \Lambda_1(1; t)\pi_1(\pi_1(\pi_0; T); t) \\
&= \Lambda_1(1; t)\frac{\pi_1(\pi_0; T)p_1(1; t)}{1 - \pi_1(\pi_0; T) + \pi_1(\pi_0; T)p_1(1; t)} \\
&= \Lambda_1(1; t)\frac{\pi_0 p_1(1; T)p_1(1; t)}{1 - \pi_0 + \pi_0 p_1(1; T)p_1(1; t)}.
\end{aligned}$$

Therefore

$$\phi(A; T, t) = (1 - \pi_0)A + \pi_0 p_1(1; T) p_1(1; t) (A - \Lambda_1(1; t)), \qquad (8.2)$$

and

$$\frac{\partial \phi(A; T, t)}{\partial T} = -\pi_0 p_1(1; T) p_1(1; t) \Lambda_1(1; T) (A - \Lambda_1(1; t)). \qquad (8.3)$$

The derivative in (8.3) is positive if $A - \Lambda_1(1; t) < 0$, i.e., if $t > \underline{T}$. □

8.6 Proof of Theorem 4.3

Since $\Psi(A, t_2^*(A, \pi_0); +\infty) < 0$,

$$A < A_2^*(\pi_0) < \hat{A}_2(\pi_0) = 0.5 \max_{t \geq 0} \Lambda_2(\pi_0; t).$$

By Lemma 2.2,

$$\Lambda_1(\pi_1(\pi_0; T_j); t) > 0.5 \Lambda_2(\pi_0, t), \forall t > T_j.$$

Therefore, for $T_j < t_2^*(A, \pi_0) < \hat{t}_2(\pi_0)$,

$$A < \hat{A}_1(\pi_1(\pi_0; T_j)) = \max_{t \geq T_j} \Lambda_1(\pi_1(\pi_0; T_j); t),$$

hence equation (4.7) has exactly two solutions $0 < t_1^*(A, \pi_1(\pi_0; T_j)) < t_{*1}(A, \pi_1(\pi_0; T_j))$. Moreover, $t_1^*(A, \pi_1(\pi_0; T_j))$ is the local maximum of $\Phi(A, \pi_0, T_j; \cdot)$; $t_{*1}(A, \pi_1(\pi_0; T_j))$ is the local minimum of $\Phi(A, \pi_0, T_j; \cdot)$ because $\Phi(A, \pi_0, T_j; \cdot)$ is increasing (respectively, decreasing) iff the LHS in (4.7) is positive (respectively, negative).

It remains to prove that $t_1^*(A, \pi_1(\pi_0; T_j))$ is the global maximum of $\Phi(A, \pi_0, T_j; \cdot)$, equivalently,

$$\int_{t_1^*}^{\infty} e^{-r(t-t_1^*)} \frac{p_1(\pi_1(\pi_0; T_j); t)}{p_1(\pi_1(\pi_0; T_j); T_j)} (A - \Lambda_1(\pi_1(\pi_0; T_j); t)) \, dt < 0. \qquad (8.4)$$

Notice that

$$\Lambda_1(\pi_1(\pi_0; T_j); T_j) = 0.5 \Lambda_2(\pi_0; T_j) < A = 0.5 \Lambda_2(\pi_0; t_2^*) < \Lambda_1(\pi_1(\pi_0; T_j); t_2^*),$$

therefore $t_1^*(A, \pi_1(\pi_0; T_j)) < t_2^*(A, \pi_0)$. Write the LHS in (8.4) as

$$\int_{t_1^*}^{t_2^*} e^{-r(t-t_1^*)} \frac{p_1(\pi_1(\pi_0; T_j); t)}{p_1(\pi_1(\pi_0; T_j); T_j)} (A - \Lambda_1(\pi_1(\pi_0; T_j); t)) \, dt$$

$$+ \frac{e^{-r(t_2^*-t_1^*)}}{p_2(\pi_0; T_j)} \int_{t_2^*}^{\infty} e^{-r(t-t_2^*)} p_2(\pi_0; T_j) \frac{p_1(\pi_1(\pi_0; T_j); t)}{p_1(\pi_1(\pi_0; T_j); T_j)} (A - \Lambda_1(\pi_1(\pi_0; T_j); t)) \, dt.$$

The first integral in the sum above is negative, because $A < \Lambda_1(\pi_1(\pi_0; T_j); t)$ for $t_1^* < t \le t_2^*$. The integrand in the second integral is $e^{-r(t-t_2^*)}\phi(A; T_j, t)$. It follows immediately from (3.21), (3.22) and (3.24) with $T_j' > T_i = t$ and $T = T_j < t$ that, for $t \in (\underline{T} \vee T, +\infty)$,

$$\phi(T, t) \le p_2(\pi_0; t)(A - 0.5\Lambda_2(\pi_0; t)). \tag{8.5}$$

On the strength of (8.5), the second integral is smaller than

$$\frac{e^{-r(t_2^*-t_1^*)}}{p_2(\pi_0; T_j)} \int_{t_2^*}^{\infty} e^{-r(t-t_2^*)} p_2(\pi_0; t)(A - 0.5\Lambda_2(\pi_0; t))\, dt$$

$$= \frac{e^{-r(t_2^*-t_1^*)} p_2(\pi_0; t_2^*)}{p_2(\pi_0; T_j)} \Psi(A, t_2^*(A, \pi_0); +\infty) < 0.$$

This proves (8.4). For any $T_j' > T_j$, $\pi_1(\pi_0; T_j') < \pi_1(\pi_0; T_j)$, hence, on the strength of (2.9), $\Lambda_1(\pi_1(\pi_0; T_j'); t) < \Lambda_1(\pi_1(\pi_0; T_j); t)$ for all $t > 0$. Therefore, $t_1^*(A, \pi_1(\pi_0; T_j')) > t_1^*(A, \pi_1(\pi_0; T_j))$. This concludes the proof of Theorem 4.3.

8.7 Proof of Lemma 5.1

Rewrite inequality (3.12) using the following steps.

(i) Use integration by parts to calculate the integral

$$\int_0^T re^{-rt} p_2(\pi_0; t)dt = 1 - e^{-rT} p_2(\pi_0; T) + \int_0^T e^{-rt} \partial_t p_2(\pi_0; t)dt$$

$$= 1 - e^{-rT} p_2(\pi_0; T) - \int_0^T e^{-rt} p_2(\pi_0; t)\Lambda_2(\pi_0; t)dt.$$

The second equality follows from the fact that

$$\partial_t p_2(\pi_0; t) = \pi_0 \partial_t p_1^2(1, t) = -2\Lambda_1(1, t)\pi_0 p_1^2(1, t)$$
$$= -2\Lambda_1(1; t)\pi_2(\pi_0; t)p_2(\pi_0; t) = \Lambda_2(\pi_0; t)p_2(\pi_0; t).$$

Hence, we can write

$$Se^{-rT} p_2(\pi_0; T) - S = -S \int_0^T e^{-rt} p_2(\pi_0; t)(r + \Lambda_2(\pi_0; t))\, dt,$$

therefore (3.12) is equivalent to

$$\sup_{T>0}\left(\int_0^T e^{-rt} p_2(\pi_0; t)((R - S)0.5\Lambda_2(\pi_0; t) - r(C + S))\, dt\right) > 0. \tag{8.6}$$

(ii) Recall the notation $\tilde{A} = r(C + S)/(R - S)$, then rewrite (8.6) as

$$\sup_{T>0} \left(\int_0^T e^{-rt} p_2(\pi_0; t) \left(0.5\Lambda_2(\pi_0; t) - \tilde{A} \right) dt \right) > 0. \qquad (8.7)$$

By definition (3.15) of $\tilde{\Psi}(\tilde{A}, 0; T)$, (8.7) is equivalent to (5.2).

8.8 Proof of Theorem 5.3

Since (8.6) holds, $\tilde{A} < \hat{A}_2(\pi_0)$. By Lemma 2.2, $0.5\Lambda_2(\pi_0; t)$ and $\Lambda_1(\pi_1(\pi_0; T_j); t)$ intersect only at $t = 0$ and $t = T_j$, therefore $\hat{A}_1(\pi_1(\pi_0, T_j)) = \max_{t \geq T_j} \Lambda_1 (\pi_1(\pi_0; T_j); t) > \tilde{A}$. Indeed, if $T_j < \hat{t}_2(\pi_0)$, then, by Lemma 2.2, $\hat{A}_1(\pi_1(\pi_0, T_j)) > \hat{A}_2(\pi_0)$, and if $\hat{t}_2(\pi_0) \leq T_j < t_{*2}(\tilde{A}, \pi_0)$, then

$$\hat{A}_1(\pi_1(\pi_0, T_j)) \geq \Lambda_1(\pi_1(\pi_0; T_j); T_j) = 0.5\Lambda_2(\pi_0; T_j) > \tilde{A}.$$

Hence equation (5.5) has two solutions

$$0 < t_1^*(\tilde{A}, \pi_1(\pi_0; T_j)) < \hat{t}_1(\pi_1(\pi_0; T_j) < t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j)).$$

Moreover, $t_1^*(\tilde{A}, \pi_1(\pi_0; T_j))$ is the local minimum of $\tilde{\Phi}(\tilde{A}, T_j; \cdot)$; and $t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$ is the local maximum of $\tilde{\Phi}(\tilde{A}, T_j; \cdot)$, because $\tilde{\Phi}(\tilde{A}, T_j; \cdot)$ is increasing (respectively, decreasing) if the LHS in (5.5) is negative (respectively, positive).

Since $T_j < t_{*2}(\tilde{A}, \pi_0)$, $\tilde{A} < 0.5\Lambda_2(\pi_0; T_j) = \Lambda_1(\pi_1(\pi_0; T_j); T_j)$. Therefore $t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j)) > T_j$. Moreover, $t_{*1} = t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j)) > t_{*2} = t_{*2}(\tilde{A}, \pi_0)$, because by Lemma 2.2, $\Lambda_1(\pi_1(\pi_0; T_j); t_{*2}) > 0.5\Lambda_2(\pi_0, t_{*2}) = \tilde{A}$. By inequality (8.5), which holds with the opposite sign in the model with conclusive breakthroughs, $\tilde{\Psi}(\tilde{A}; T_j, t_{*2}) \leq \tilde{\Phi}(\tilde{A}, T_j; t_{*2})$. Since (8.6) holds, we also have

$$0 < \tilde{\Psi}(\tilde{A}; 0, t_{*2}) \leq \tilde{\Psi}(\tilde{A}; T_j, t_{*2}) \leq \tilde{\Phi}(\tilde{A}, T_j; t_{*2}).$$

By definition of a maximum, $\tilde{\Phi}(\tilde{A}, T_j; t_{*1}) \geq \tilde{\Phi}(\tilde{A}, T_j; t_{*2}) > 0$, hence $t_{*1} = t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$ is the global maximum of $\tilde{\Phi}(\tilde{A}, T_j; \cdot)$. For any $T_j' > T_j$, $\pi_1(\pi_0; T_j') < \pi_1(\pi_0; T_j)$, hence, on the strength of (2.9), $\Lambda_1(\pi_1(\pi_0; T_j'); t) < \Lambda_1(\pi_1(\pi_0; T_j); t)$ for all $t > 0$. Therefore, $t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j')) < t_{*1}(\tilde{A}, \pi_1(\pi_0; T_j))$.

8.9 Proof that $t_{*2}(\tilde{A}, \pi_0)$ is not a point of maximum the social planner's function in the good news model

It suffices to show that the second-order derivative $\partial_{TT}^2 S(\tilde{A}; t_{*2}-) > 0$. To this end, will show that $\partial_{TT}^2 S(\tilde{A}; t_{*2}-)$ is given by (5.13). Using (8.3) for the partial derivative under the integral in (5.12), we rewrite (5.12) as

$$\frac{\partial SP(\tilde{A}; T)}{\partial T} = \mathrm{e}^{-rT} p_2(\pi_0; T) \left(0.5\Lambda_2(\pi_0; T) - \tilde{A}\right)$$
$$+ \pi_0 p_1(1; T)\Lambda_1(1, T) \int_T^{T_f(T)} \mathrm{e}^{-rt} p_1(1; t)(A - \Lambda_1(1; t))\mathrm{d}t. \tag{8.8}$$

Consider $T = t_{*2} - \Delta t$. Then the first term in (8.8) is

$$\mathrm{e}^{-r(t_{*2} - \Delta t)} p_2(\pi_0; (t_{*2} - \Delta t)) \left(0.5\Lambda_2(\pi_0; (t_{*2} - \Delta t)) - \tilde{A}\right)$$
$$= -\mathrm{e}^{-rt_{*2}} p_2(\pi_0; t_{*2}) 0.5\partial_t \Lambda_2(\pi_0; t_{*2})\Delta t + o(\Delta t). \tag{8.9}$$

The second term in (8.8) is

$$\pi_0 p_1(1; t_{*2} - \Delta t)\Lambda_1(1, t_{*2} - \Delta t) \int_{t_{*2} - \Delta t}^{T_f(t_{*2} - \Delta t)} \mathrm{e}^{-rt} p_1(1; t)(A - \Lambda_1(1; t))\mathrm{d}t$$
$$= \mathrm{e}^{-rt_{*2}} \pi_0 p_1(1; t_{*2})^2 \Lambda_1(1, t_{*2})(A - \Lambda_1(1, t_{*2})) \left(1 - \frac{\partial T_f(t_{*2}-)}{\partial T}\right) \Delta t + o(\Delta t).$$

By (2.9), $\pi_0 p_1(1; t_{*2})^2 \Lambda_1(1, t_{*2}) = 0.5\Lambda_2(\pi_0, t_{*2}) p_2(\pi_0, t_{*2})$, hence the second term in (8.8) is equal to

$$\mathrm{e}^{-rt_{*2}} p_2(\pi_0, t_{*2}) 0.5\Lambda_2(\pi_0, t_{*2})(A - \Lambda_1(1, t_{*2})) \left(1 - \frac{\partial T_f(t_{*2}-)}{\partial T}\right) \Delta t + o(\Delta t). \tag{8.10}$$

To find $\partial T_f(t_{*2} - 0)/\partial T$, we use (5.5) and the implicit function theorem to derive

$$\frac{\partial T_f(T)}{\partial T} = -\frac{\partial \Lambda_1(\pi_1(\pi_1; T); T_f)/\partial T}{\partial_t \Lambda_1(\pi_1(\pi_1; T); T_f)}. \tag{8.11}$$

On the strength of (2.9) and (2.6), the numerator in the last fraction is

$$\frac{\partial \Lambda_1(\pi_1(\pi_1; T); T_f)}{\partial T} = -\Lambda_1(1; T)\Lambda_1(1; T_f)\pi_1(\pi_1(\pi_0; T); T_f)$$
$$(1 - \pi_1(\pi_1(\pi_0; T); T_f)) < 0.$$

As $T \uparrow t_{*2}$, the above expression tends to $-\Lambda_1(1; t_{*2})^2 \pi_2(\pi_0; t_{*2})(1 - \pi_2(\pi_0; t_{*2}))$. By (2.11), the denominator in (8.11) is

$$\partial_t \Lambda_1(\pi_1(\pi_1; T); T_f) = \pi_1(\pi_1(\pi_0; T); T_f)$$
$$\left[\partial_t \Lambda_1(1; T_f) - \Lambda_1(1; T_f)^2(1 - \pi_1(\pi_1(\pi_0; T); T_f))\right].$$

As $T \uparrow t_{*2}$, the last expression tends to

$$\pi_2(\pi_0; t_{*2}) \left[\partial_t \Lambda_1(1; t_{*2}) - \Lambda_1(1; t_{*2})^2 (1 - \pi_2(\pi_0; t_{*2})) \right].$$

Hence

$$
\begin{aligned}
1 - \frac{\partial T_f(t_{*2}-)}{\partial T} &= 1 - \frac{\Lambda_1(1; t_{*2})\pi_2(\pi_0; t_{*2})(1 - \pi_2(\pi_0; t_{*2}))}{\pi_2(\pi_0; t_{*2}) \left[\partial_t \Lambda_1(1; t_{*2}) - \Lambda_1(1; t_{*2})^2 (1 - \pi_2(\pi_0; t_{*2})) \right]} \\
&= \frac{\pi_2(\pi_0; t_{*2}) \left[\partial_t \Lambda_1(1; t_{*2}) - 2\Lambda_1(1; t_{*2})^2 (1 - \pi_2(\pi_0; t_{*2})) \right]}{\pi_2(\pi_0; t_{*2}) \left[\partial_t \Lambda_1(1; t_{*2}) - \Lambda_1(1; t_{*2})^2 (1 - \pi_2(\pi_0; t_{*2})) \right]} \\
&= \frac{0.5\partial_t \Lambda_2(\pi_0; t_{*2})}{\pi_2(\pi_0; t_{*2}) \left[\partial_t \Lambda_1(1; t_{*2}) - \Lambda_1(1; t_{*2})^2 (1 - \pi_2(\pi_0; t_{*2})) \right]}.
\end{aligned}
\tag{8.12}
$$

The last equality holds on the strength of (2.11).

Using (8.8), (8.9), (8.10), and (8.12), we derive

$$
\frac{\partial SP(\tilde{A}; t_{*2} - \Delta t)}{\partial T} = \Delta t e^{-rt_{*2}} p_2(\pi_0; t_{*2}) 0.5\partial_t \Lambda_2(\pi_0; t_{*2})
$$
$$
\times \left(-1 + \frac{0.5\Lambda_2(\pi_0, t_{*2})(A - \Lambda_1(1, t_{*2}))}{\pi_2(\pi_0; t_{*2}) \left[\partial_t \Lambda_1(1; t_{*2}) - \Lambda_1(1; t_{*2})^2 (1 - \pi_2(\pi_0; t_{*2})) \right]} \right)
$$
$$
+ o(\Delta t).
$$

It remains to replace A with $0.5\Lambda_2(\pi_0, t_{*2})$ and use (2.9) for $0.5\Lambda_2(\pi_0, t_{*2})$ in the term in parenthesis in order to obtain

$$
\frac{\partial SP(\tilde{A}; t_{*2} - \Delta t)}{\partial T} = \Delta t \frac{-e^{-rt_{*2}} p_2(\pi_0; t_{*2}) 0.5\partial_t \Lambda_2(\pi_0; t_{*2}) \pi_2(\pi_0; t_{*2}) \partial_t \Lambda_1(1; t_{*2})}{\partial_t \Lambda_1(\pi_1(\pi_0; t_{*2}); t_{*2})}
$$
$$
+ o(\Delta t).
$$

Dividing both sides of the last equation by $-\Delta t$ and passing to the limit $\Delta t \to 0$, we obtain (5.13).

References

Amir, R.: Continuous stochastic games of capital accumulation with convex transitions. Games Econ. Behav. **15**, 111–131 (1996a)

Amir, R.: Cournot oligopoly and the theory of supermodular games. Games Econ. Behav. **15**, 132–148 (1996b)

Amir, R., Lazzati, N.: Network effects, market structure and industry performance. J. Econ. Theory **146**, 2389–2419 (2011)

Balbus, Ł., Reffett, K., Woźny, Ł.: A constructive study of markov equilibria in stochastic games with strategic complementarities. J. Econ. Theory **150**, 815–840 (2014)

Bergemann, D., Välimäki, J.: Learning and strategic pricing. Econometrica **64**, 1125–1149 (1996)

Bergemann, D., Välimäki, J.: Experimentation in markets. Rev. Econ. Stud. **67**, 213–234 (2000)

Bergemann, D., Välimäki, J.: Dynamic price competition. J. Econ. Theory **127**, 232–263 (2006)

Bergemann, D., Välimäki, J.: Bandit problems. In: Durlauf, S.N., Blume, L.E. (eds.) The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke (2008)

Bolton, P., Harris, C.: Strategic experimentation. Econometrica **67**, 349–374 (1999)

Boyarchenko, S.: Strategic experimentation with humped bandits. Working paper. Available at SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3174107 (2018). Accessed 16 May 2019

Boyarchenko, S.I., Levendorksiĭ, S.Z.: Preemption games under Lévy uncertainty. Games Econ. Behav. **88**, 354–380 (2014)

Curtat, L.O.: Markov equilibria of stochastic games with complementarities. Games Econ. Behav. **17**, 177–199 (1996)

Das, K., Klein, N., Schmid, K.: Strategic experimentation with asymmetric players. Econ. Theory. https://doi.org/10.1007/s00199-019-01193-9 (2019)

Decamps, J.-P., Mariotti, T.: Investment timing and learning externalities. J. Econ. Theory **118**, 80–102 (2004)

Dutta, P.K., Rustichini, A.: A theory of stopping time games with applications to product innovations and asset sales. Econ. Theory **3**, 743–763 (1993). https://doi.org/10.1007/BF01210269

Fudenberg, D., Tirole, J.: Preemption and rent equalization in the adoption of new technology. Rev. Econ. Stud. **52**, 383–401 (1985)

Halac, M., Kartik, N., Liu, Q.: Contests for experimentation. J. Econ. Theory **125**, 1523–1569 (2017)

Heidhues, P., Rady, S., Strack, P.: Strategic experimentation with private payoffs. J. Econ. Theory **159**, 531–551 (2015)

Hörner, J., Skrzypacz, A.: Learning, experimentation and information design. Working Paper, Stanford University (2016)

Hörner, J., Klein, N.A., Rady, S.: Strongly symmetric equilibria in bandit games. Cowles Foundation Discussion Paper No. 1056. Available at SSRN. http://ssrn.com/abstract=2482335 (2014). Accessed 16 May 2019

Keller, G., Rady, S.: Strategic experimentation with Poisson bandits. Theor. Econ. **5**, 275–311 (2010)

Keller, G., Rady, S.: Breakdowns. Theor. Econ. **10**, 175–202 (2015)

Keller, G., Rady, S., Cripps, M.: Strategic experimentation with exponential bandits. Econometrica **73**, 39–68 (2005)

Khan, U., Stinchcombe, M.B.: The virtues of hesitation: optimal timing in a non-stationary world. Am. Econ. Rev. **105**, 1147–1176 (2015)

Klein, N., Rady, S.: Negatively correlated bandits. Rev. Econ. Stud. **78**, 693–732 (2008)

Kolmogorov, A.N., Fomin, S.V.: Introductory real analysis (Translated and edited by Silverman, R.A). Dover Books in Mathematics (1975)

Laraki, R., Solan, E., Vieille, N.: Continuous-time games of timing. J. Econ. Theory **120**, 206–238 (2005)

Laussel, D., Resende, J.: Dynamic price competition in aftermarkets with network effects. J. Math. Econ. **50**, 106–118 (2014)

Marlats, C., Ménager, L.: Strategic observation with exponential bandits. Working Paper (2018)

Pawlina, G., Kort, P.M.: Real options in asymmetric duopoly: who benefits from your comparative disadvantage? J. Econ. Manag. Strategy **15**, 1–35 (2006)

Riedel, F., Steg, J.H.: Subgame-perfect equilibria in stochastic timing games. J. Math. Econ. **72**, 36–50 (2017)

Rosenberg, D., Salomon, A., Vieille, N.: On games of strategic experimentation. Games Econ. Behav. **82**, 31–51 (2013)

Thijssen, J.J.J., Huisman, K.J.M., Kort, P.M.: The effects of information of strategic investment and welfare. Econ. Theory **28**, 399–424 (2006). https://doi.org/10.1007/s00199-005-0628-3

Thijssen, J.J.J., Huisman, K.J.M., Kort, P.M.: Symmetric equilibrium strategies in game theoretic real option models. J. Math. Econ. **48**, 219–225 (2012)

Vives, X.: Oligopoly Pricing: Old Ideas and New Tools. MIT Press, Cambridge (1999)

Vives, X.: Complementarities and games: new developments. J. Econ. Lit. **43**, 37–479 (2005)

Vives, X.: Strategic complementarity in multi-stage games. Econ. Theory **40**, 151–171 (2009). https://doi.org/10.1007/s00199-008-0354-8