Social Learning under Negative Correlation in an Exit Game*

Yifan Dai† and Kyungmin Kim‡

November 2015

Abstract

We consider a two-player exit game in which each player faces a one-armed bandit problem and the two players' types are negatively correlated. We provide a closed-form characterization of the unique (perfect Bayesian) equilibrium of the game. We show that, in stark contrast to the case of positive correlation, the players exit the game at an increasing rate over time and one player exits for sure before a deterministic time.

JEL Classification Numbers: C73, D82, D83.

Keywords: Social learning; experimentation; negative correlation.

1 Introduction

Learning comes in two ways, private and social. In other words, we learn not only from our own experiences, but also from others' behavior. Furthermore, the two forms of learning are intertwined, because others also learn from our behavior. Models of strategic experimentation, in which multiple players simultaneously engage in learning, offer a natural framework to study how private learning and social learning interact each other and what their economic consequences are.

Most papers in the literature study the case of positive correlation in which good news to a player is also good news to other players (more precisely, if a player's type is good, then the other player's type is also, or more likely to be, good). Bolton and Harris (1999), Keller, Rady and Cripps (2005), and Keller and Rady (2010) consider the case of perfect social learning in which both

*We thank Rabah Amir, Hantaek Bae and Yingni Guo for various helpful comments.
†University of Iowa. Contact: yifan-dai-1@uiowa.edu
‡University of Iowa. Contact: kyungmin-kim@uiowa.edu

actions and payoffs are observable to other players.[1] Rosenberg, Solan and Vieille (2007) and Murto and Välimäki (2011) consider the case of imperfect social learning in which only actions are observable. Specifically, they analyze an exit game in which each player knows only whether the other players stay in the game or not.[2]

Negative correlation, with which good news to a player is bad news to the other player, is equally plausible, but has received disproportionally less attention: to our knowledge, Klein and Rady (2011) is the only relevant contribution. The form of correlation is not crucial for social learning per se. In particular, learning is always valuable, and thus free-riding incentives arise whether correlation is positive or negative. However, the correlation structure determines the content of social learning and, therefore, dictates the equilibrium dynamics. Klein and Rady (2011) demonstrate this in the model with perfect social learning. In their baseline model (with two players and perfect negative correlation), if both players engage in experimentation, then the players' beliefs, conditional on no success, stay constant, which has significant implications for the equilibrium structure and efficiency.[3]

We consider negative correlation in the context of an exit game. Specifically, each player decides whether to stay in the game or not (exit). They can exit at any point in the game, but exit is irreversible.[4] Each player is either good or bad. If a player is good, then he receives lump-sum rewards at a positive Poisson rate. The other player is bad and never receives a lump-sum reward. Each player prefers staying if his type is good, but prefers exiting immediately if his type is bad. This model can be interpreted as an imperfect-learning counterpart to Klein and Rady (2011), or as a negative-correlation counterpart to Murto and Välimäki (2011).

We provide a closed-form characterization of the unique (perfect Bayesian) equilibrium of the model and contrast it to the unique equilibrium under positive correlation. In both cases, the equilibrium features two phases. In the first phase, no player exits, and thus social learning does not occur. Conditional on no success, a player becomes more pessimistic solely based on his private learning. Once the players become sufficiently pessimistic, the second phase begins, in which the players exit at a positive rate, and thus a player learns not only from his own experience, but also

[1] They adopt different learning processes. Bolton and Harris (1999) consider a Brownian learning model, Keller, Rady and Cripps (2005) employ an exponential learning model (in which Poisson signals arrive at a positive rate only at the good state), and Keller and Rady (2010) examine a Poisson learning model (in which Poisson signals arrive at different rates at different states).

[2] Imperfect social learning raises a non-trivial inference problem regarding the types of the remaining players. Rosenberg, Solan and Vieille (2007) focus on perfect positive correlation, but incorporate a more general signal structure, while Murto and Välimäki (2011) consider a simple signal structure, but allow for imperfect positive correlation.

[3] Most notably, there exits an equilibrium in which each player plays a simple cutoff strategy (experimenting if and only if the probability that his type is good exceeds a certain threshold), and the resulting equilibrium is efficient for a range of parameter values. None of these results holds in the case of positive correlation (see Keller, Rady and Cripps, 2005).

[4] In our model, the player who exits first has an incentive to re-enter the game once the other player also exits. We provide a brief discussion on re-entry in Section 6.3.

from the other player's behavior. In both cases, the second phase ends upon a player's exit.

One clear difference between positive correlation and negative correlation lies in the behavior at the end of the second phase. With positive correlation, a player's exit triggers the other player's exit: a player's exit reveals that he has not succeeded yet, which makes the other player more pessimistic and, therefore, willing to exit. With negative correlation, the same news is good news to the other player, who will then revise up his belief and stay longer in the game.

More importantly, the correlation structure affects the way private learning and social learning interact and, therefore, the resulting equilibrium dynamics in the second phase. For the players to remain indifferent between staying and exiting in the second phase (which is necessary for them to exit at a positive rate), the benefit of social learning must compensate increasing pessimism due to private learning. With positive correlation, the benefit comes in the form of positive information: a player is more likely to stay in the game when he is good than when he is bad. Therefore, the fact that a player stays in the game allows the other player to overcome his own pessimism and be willing to stay in the game. Formally, this translates into the players' beliefs, conditional on no success and no exit by the other player, staying constant and the second phase stretching without a limit. With negative correlation, the corresponding benefit comes in the form of an increasing amount (speed) of social learning. A player's staying is now bad news to the other player, and thus social learning makes the players more pessimistic. This implies that the players, conditional on no success and no exit by the other player, necessarily become more pessimistic over time. This growing pessimism can be compensated only through even more amount of social learning, which translates into the players, conditional on no success, exiting the game at an increasing rate over time. Growing pessimism and increasing (negative) social learning reinforce each other. As a result, under negative correlation, in stark contrast to the positive-correlation case, the second phase necesssarily ends by a finite deterministic time (by which one player exits for sure and, therefore, the players' conditional beliefs converge to 0).

The rest of the paper is organized as follows. We introduce the model in Section 2 and present two benchmark models, the single-player problem and the positive-correlation case, in Section 3. We analyze the symmetric case (in which the two players are ex ante identical) in Section 4 and the asymmetric case in Section 5. We discuss three relevant extensions in Section 6.

2 The Model

We set up the model in continuous time. Time starts from 0 and is indexed by $t \in \mathcal{R}_+$. There are two players, player 1 and player 2. Per each time unit $[t, t+dt)$, the players first decide whether to stay in the game or exit the game. Staying is costly: each player incurs flow cost c per unit time. Exit is costless but irreversible: once a player exits, he cannot reenter the game.

Each player is either good or bad. If player i is good, then he constantly receives lump-sum payoff $v(> 0)$ at a Poisson rate λ. If player i is bad, then he never receives a lump-sum payoff.[5] The payoff a player receives when he exits the game is normalized to 0. In order to avoid triviality, we assume that a player strictly prefers staying in the game to exiting if his type is good, while the opposite is true if his type is bad. Formally, we assume that $\lambda v > c > 0$.

The players are initially uncertain about their types. Denote by p_i the prior probability that player i's type is good. The players' types are perfectly negatively correlated: if player i is good, then player j is bad. It is necessary that $p_1 + p_2 = 1$. This information structure is common knowledge between the players.

Each player's action is observable to the other player, but his payoff is not. This means that when player j stays in the game, player i is not sure whether player j has already succeeded or not. Once a player receives a lump-sum payoff, it is a dominant strategy for him to stay in the game. Therefore, player j's exit is good news to player i, as it reveals that player j has not succeeded, which is more likely when player j is bad (i.e., player i is good).

Denote by 0 "stay" and by 1 "exit". An action profile at time t is then a vector $a^t = (a_1^t, a_2^t) \in \{0, 1\}^2$. The public history of the game is then a sequence of action profiles. Let H^t denote the set of all histories until time t, and $H^0 \equiv \emptyset$. Finally, define the set of all histories $H \equiv \cup_t H^t$.

Each player's private history consists of the public history and his past realized payoffs. Since the optimal strategy of a player who has ever received payoff v is straightforward, it suffices to consider the private histories up to which each player has not received payoff v. This implies that each player's strategy can be defined as a function of the public history. Formally, player i's pure strategy, conditional on no success, can be defined as a function $s_i : H \to \{0, 1\}$, where $s_i(h^t)$ represents player i's exit decision following history h^t. Since exit is irreversible, player i's strategy s_i is admissible only when if $s_i(h^t) = 1$, then $s_i(h^s) = 1$ for any history h^s following h^t. Player i's mixed strategy is a probability distribution over the set of player i's admissible pure strategies.

Each player maximizes his expected discounted sum of payoffs and is risk neutral. We study perfect Bayesian equilibrium of this game: for both $i = 1, 2$, player i's strategy is a best response to player j's strategy after any history $h^t \in H$, and the players' beliefs after each history are obtained by Bayes' rule whenever possible.

[5] Notice that once a player receives payoff v, he becomes sure that his type is good. This is a common simplifying assumption in the literature (see, e.g., Keller, Rady and Cripps, 2005; Klein and Rady, 2011; Murto and Välimäki, 2011; Bonatti and Hörner, 2011). It is well-known that if this assumption is relaxed (i.e., a player may receive payoff v even if her type is bad), then the analysis becomes significantly more complicated (see, among others, Keller and Rady, 2010).

3 Two Benchmarks

This section provides the results for two benchmark models, one without social learning (i.e., the single-player problem) and the other with positive correlation.

3.1 No Social Learning

We first consider the case where a player does not observe the other player's action. This case is formally identical to the single-player experimentation problem, which is familiar in the literature. In particular, the following result is a straightforward modification of Proposition 3.1 in Keller, Rady and Cripps (2005).

Proposition 1 *In the absence of social learning, each player stays in the game if and only if he assigns a greater probability than* p^* *to the event that his type is good, where*

$$p^* \equiv \frac{c}{\lambda\left(v + \frac{\lambda v - c}{r}\right)}.$$

Conditional on no success, his belief (the probability that he is good) decreases according to $\dot{p}(t) = -\lambda p(t)(1-p(t))$*. His expected payoff as a function of his belief is equal to*

$$V(p) = \begin{cases} 0, & \textit{if } p \leq p^*, \\ \frac{p\lambda v - c}{r} + \frac{c - p^*\lambda v}{r}\frac{1-p}{1-p^*}\left(\frac{1-p}{p}\frac{p^*}{1-p^*}\right)^{r/\lambda}, & \textit{if } p > p^*. \end{cases}$$

The result implies that in the absence of social learning, each player's optimal strategy takes a simple form: he stays in the game only until his belief reaches p^*. The length of time player i stays in the game, denoted by t_i^*, can be explicitly calculated as follows: if $p_i \leq p^*$, then $t_i^* = 0$ (i.e., immediate exit). Otherwise, his belief must be p^* at time t_i^*. Therefore,

$$p^* = \frac{p_i e^{-\lambda t_i^*}}{p_i e^{-\lambda t_i^*} + 1 - p_i} \Rightarrow t_i^* = -\frac{1}{\lambda} log\left(\frac{1-p_i}{p_i}\frac{p^*}{1-p^*}\right). \quad (1)$$

For later use, define $t^* \equiv \min\{t_1^*, t_2^*\}$.

3.2 Positive Correlation

Now we consider the case where the players' types are positively correlated: if player i is good (bad), then player j is also good (bad). This case has been extensively studied in the literature. In particular, our model is a special case of Murto and Välimäki (2011), with two players and perfect correlation. Note that the players should have the same prior beliefs in this case (i.e., $p_1 = p_2$).

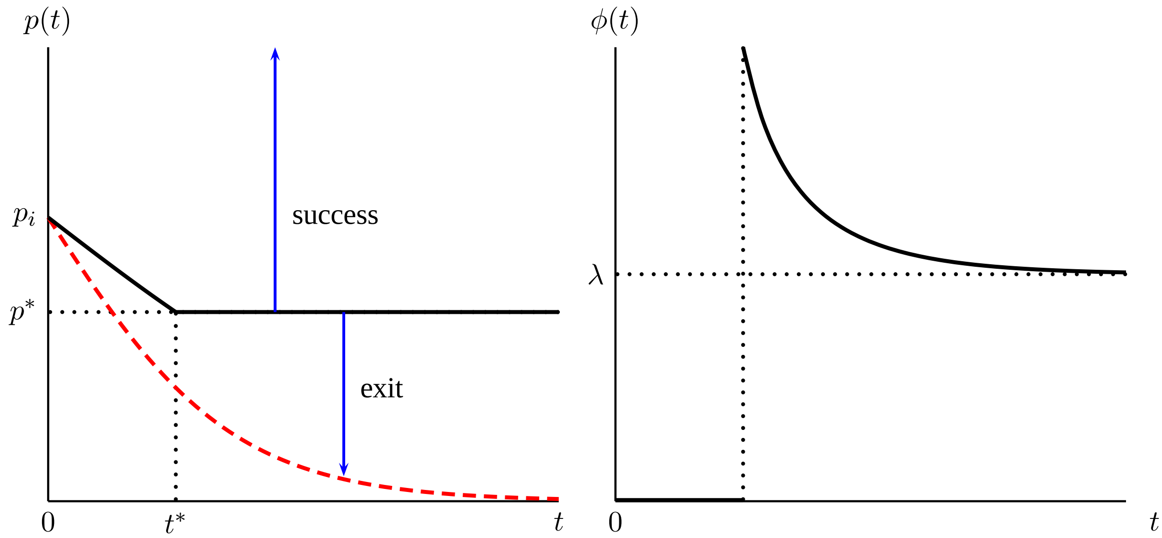

Figure 1: Each player's belief $p(t)$ (left) and exit rate $\phi(t)$ (right) at time t, conditional on no success and no exit by the other player. The dashed line in the left panel represents the players' beliefs when they know that both of them have not succeeded by time t (i.e., $p_i e^{-2\lambda t}/(p_i e^{-2\lambda t} + 1 - p_i)$).

The following proposition provides a closed-form characterization of the unique equilibrium.

Proposition 2 *In the model with perfect positive correlation, there exists a unique equilibrium. Until time t^*, no player exits and the players' conditional beliefs $p(t)$ decrease according to $\dot{p}(t) = -\lambda p(t)(1 - p(t))$. After time t^*, each player, conditional on no success, exits at a decreasing rate:*

$$\phi(t) = \frac{\lambda p^*(1 - p_i)e^{2\lambda t}}{((1 - p_i)e^{2\lambda t} + p_i)p^* - p_i},$$

and the players' conditional beliefs $p(t)$ stay constant at p^*. If one player exits, then the other player follows immediately.

Proof. See the appendix. ∎

Figure 1 illustrates the resulting equilibrium dynamics. Until time t^*, no player exists. Therefore, each player learns only from his own experience (failure), updating his belief as in the single-player problem. Once the players' beliefs reach p^*, they randomize between staying and exiting at a well-defined rate: if player i stays for sure, then player j does not learn from player i's action and, therefore, exits immediately, unless he has already succeeded and, therefore, knows that his type is good. This delivers a significant amount of information to player i, who will then follow player j's action: conditional on no success, player i exits if player j exits and stays if player j stays. This, in turn, provides a lot of information for player j, which deters player j's exit in the first place, unraveling the given equilibrium structure. Following the same logic, it is also clear that no player exits with a positive probability at each point in time.

The players' beliefs $p(t)$, conditional on no success and no exit, stay constant once they reach p^* (see the left panel of Figure 1). There are two opposing effects. On the one hand, player i's own failure pushes down his belief: without social learning, his belief would keep decreasing as in the single-player problem. On the other hand, player j's staying is good news to player i: player j is more likely to stay when he is good than when he is bad. In equilibrium, these two effects are balanced, and thus $p(t)$ stays constant after t^*. The equilibrium exit rate $\phi(t)$ is strictly decreasing over time (see the right panel of Figure 1). This is because each player is more likely to know his type and, therefore, staying becomes an increasingly better indicator of the good type as t increases. The exit rate $\phi(t)$ converges to λ. This is because in the limit, player j knows his type for sure, and thus his exiting at rate λ (when his type is bad) suffices to compensate player i's own failure and restore player i's belief back to p^*.

4 Symmetric Negative Correlation

We now study the main model with negative correlation. We first consider the symmetric case where $p_1 = p_2 = 1/2$, which is directly comparable to the case of positive correlation.

Evolution of the players' conditional beliefs. We begin by characterizing how the players' beliefs evolve over time. Specifically, we derive how each player' belief $p(t)$, conditional on no success and no exit, changes over time when the other player's exit strategy is given by $\phi(t)$.

Conditional on player i being good, the probability that player i does not succeed until time t is equal to $e^{-\lambda t}$. Player j never succeeds and the probability that player j does not exit by time t is given by $e^{-\int_0^t \phi(y)dy}$.

Conditional on player j being good, player i never succeeds. Player j stays for sure if he has succeeded, but might have exited if success arrives rather late. To formally derive the relevant probability, let x denote player j's first success time, which is exponentially distributed with parameter λ. Player j stays until time t as long as he has not exited by time $\min\{x,t\}$, whose probability is equal to $e^{-\int_0^{\min\{x,t\}} \phi(y)dy}$. It follows that the probability that player j stays until time t is equal to $\int_0^\infty e^{-\int_0^{\min\{x,t\}} \phi(y)dy} d(1-e^{-\lambda x})$.

Combining the above two cases, by Bayes' rule, the probability that player i assigns to the event that his type is good at time t, conditional on no success and no exit, is given by

$$p(t) = \frac{p_i e^{-\lambda t} e^{-\int_0^t \phi(y)dy}}{p_i e^{-\lambda t} e^{-\int_0^t \phi(y)dy} + p_j \int_0^\infty e^{-\int_0^{\min\{x,t\}} \phi(y)dy} d(1-e^{-\lambda x})}.$$

Arranging the terms and using the fact that $p_1 = p_2 = 1/2$, the expression shrinks to

$$p(t) = \frac{1}{2 + \int_0^t e^{\int_x^t (\lambda + \phi(y))dy} \lambda dx}, \quad (2)$$

which implies

$$\dot{p}(t) = -\lambda p(t)(1 - p(t)) - \phi(t)p(t)(1 - 2p(t)). \quad (3)$$

The players' conditional beliefs $p(t)$ always decrease over time. There are two reasons. First, each player learns from his own failure. The first term in equation (3) represents this effect. Second, each player learns from the other player's action. Since player j is more likely to stay when his type is good, player j's staying is bad news to player i, further pushing down player i's belief. This effect is captured in the second term in equation (3).

The players' conditional value function. In the symmetric equilibrium, each player's expected payoff as a function of his belief $p(t)$ is identical to that without social learning in Proposition 1.[6] Intuitively, player i learns from player j's action only when player j may exit (i.e., $\phi(t) > 0$). However, if $\phi(t) > 0$, then exit is an optimal strategy for player i. In other words, social learning occurs only when it is irrelevant to the players' expected payoffs.

Together with the fact that $p(t)$ is strictly decreasing, this significantly simplifies the analysis. No player exits until his belief reaches p^*. This implies that $p(t)$ reaches p^* at time t^*. If $t > t^*$, then each player's expected payoff, conditional on no success and no exit, stays constant at 0: recall that $V(p) = 0$ if $p \leq p^*$ in Proposition 1, and $p(t)$ always decreases.

To utilize the fact that the players' conditional expected payoffs remain equal to 0 after time t^*, we calculate the probability that player i assigns to the event that player j has not succeeded, conditional on his staying until time $t(> t^*)$. If player i is good, then player j is bad and, therefore, has not succeeded for sure. If player j is good, then the probability that he stays and has succeeded before time t is equal to $\int_0^t e^{-\int_0^x \phi(y)dy} d(1 - e^{-\lambda x})$, while the probability that he stays but has not succeeded until time t is equal to $e^{-\int_0^t (\lambda + \phi(y))dy}$. Therefore, the total probability that player j stays but has not succeeded by time t is equal to

$$p(t) + (1 - p(t)) \frac{e^{-\int_0^t (\lambda + \phi(y))dy}}{e^{-\int_0^t (\lambda + \phi(y))dy} + \int_0^t e^{-\int_0^x \phi(y)dy} d(1 - e^{-\lambda x})} = p(t) + (1 - p(t)) \frac{1}{1 + \int_0^t e^{\int_x^t (\lambda + \phi(y))dy} \lambda dx}.$$

Applying equation (2), the probability simplifies to $2p(t)$.

[6] Murto and Välimäki (2011) prove this result in the model with positive correlation (Lemma 1 in their paper). Despite the difference in the correlation structure, their argument applies unchanged to our model with negative correlation. What is crucial for the result is the irreversibility of exit. The result does not hold if the players can reenter the game.

Combining all the results so far, the following equation holds whenever $t > t^*$:

$$\begin{aligned} 0 = V(p(t)) &= -cdt + 2p(t)\phi(t)dt \cdot V(0.5) \\ &+(1-2p(t)\phi(t)dt)\left(p'(t)\lambda dt \cdot (v + e^{-rdt}V(1)) + (1-p'(t)\lambda dt)e^{-rdt}V(p(t+dt))\right), \end{aligned}$$

where

$$p'(t) = \frac{p(t)(1-\phi(t)dt)}{p(t)(1-\phi(t)dt) + (1-p(t))\frac{(1-\phi(t)dt)+\int_0^t e^{\int_x^t(\lambda+\phi(y))dy}\lambda dx}{1+\int_0^t e^{\int_x^t(\lambda+\phi(y))dy}\lambda dx}}.$$

The left-hand side is a player's expected payoff when he exits the game, while the right-hand side is his expected payoff when staying. The right-hand side consists of four terms: the first term is the flow cost of staying. The second term represents the possibility that the other player exits, in which case the player's belief jumps to $1/2$ and he faces the same problem as in Proposition 1. The last two terms correspond to the case where the other player stays in the game. In that case, the player's belief updates to $p'(t)$, the player succeeds with probability $p'(t)\lambda dt$, and his continuation payoff becomes either $V(1)$ or $V(p(t+dt))$, depending on whether he succeeds or not.

Arranging the terms, the expression simplifies to

$$c = 2p(t)\phi(t)V(0.5) + p(t)\lambda(v + V(1)) \Leftrightarrow \phi(t) = \frac{\frac{c}{p(t)} - \lambda(u+V(1))}{2V(0.5)}. \tag{4}$$

Intuitively, the left-hand side is the marginal cost of staying an instant longer, while the right-hand side is the corresponding marginal benefit. The latter comes from the fact that the other player may exit, which occurs at rate $\phi(t)$ in case the other player has not succeeded yet (whose probability is equal to $2p(t)$), or the player may succeed, which occurs at rate λ when his type is good (whose probability is equal to $p(t)$). The value of $\phi(t)$ is well-defined whenever $t \geq t^*$, because $c = p^*\lambda(u + V(1)) \geq p(t)\lambda(u + V(1))$. It is strictly decreasing in $p(t)$ and, therefore, strictly increasing in t. Intuitively, player i becomes increasingly pessimistic as he continues to fail, while player j does not exit. Since this increases player i's incentive to exit, player j must exit at an increasing rate, so as to provide a stronger incentive for player i to stay.

Equilibrium characterization. Combining equations (3) and (4) yields the following non-linear differential equation for $p(t)$:

$$\dot{p}(t) = -\lambda p(t)(1-p(t)) - \frac{c - p(t)\lambda(v + V(1))}{2V(0.5)}(1-2p(t)).$$

Arranging the terms leads to

$$\dot{p}(t) = -\frac{c}{2V(0.5)} + \frac{2c + \lambda(v + V(1) - 2V(0.5))}{2V(0.5)} p(t) - \frac{\lambda(v + V(1)) - \lambda V(0.5)}{V(0.5)} p^2(t). \quad (5)$$

This is a quadratic first-order differential equation, known as a Riccati equation, with constant coefficients, and admits a closed-form solution, as reported in the following lemma for convenience.[7]

Lemma 1 *Suppose $p(t)$ is a deterministic function of time t, $p(t^*) = p^*$, and satisfies*

$$\dot{p}(t) = A + Bp(t) + Cp^2(t),$$

where A, B, and C are constant real numbers. The solution to the differential equation is[8]

$$p(t) = -\frac{k_2}{C} \frac{\frac{k_1}{k_2}\frac{k_2+Cp^*}{k_1+Cp^*} e^{(k_1-k_2)(t-t^*)} - 1}{\frac{k_2+Cp^*}{k_1+Cp^*} e^{(k_1-k_2)(t-t^*)} - 1}, \quad (6)$$

where

$$k_1 = \frac{B + \sqrt{B^2 - 4AC}}{2}, \text{ and } k_2 = \frac{B - \sqrt{B^2 - 4AC}}{2}.$$

Applying Lemma 1 to equation (5) and the resulting solution $p(t)$ to equation (4), we obtain the following result.

Proposition 3 *In the symmetric case with negative correlation, there exits a unique equilibrium. Until time t^*, no player exits and the players' conditional beliefs $p(t)$ decrease according to $\dot{p}(t) = -\lambda p(t)(1-p(t))$. After time t^*, the players' conditional beliefs $p(t)$ decrease according to equation* (5) *(whose solution can be obtained through Lemma 1), and each player exits at rate $\phi(t)$, as given in equation* (4). *If one player exits, then the other player updates his belief to $1/2$ and behaves as described in Proposition 1.*

Proof. See the appendix. ■

[7]This equation arises in various contexts in macroeconomics and finance. See, e.g., Ljungqvist and Sargent (2004) and Nawalkha, Soto and Beliaeva (2007). In most cases, the structure is either exogenously imposed or obtained as a solution to a linear quadratic dynamic programming problem (see Chapter 5 in Ljungqvist and Sargent, 2004). We are unaware of any other model in which a Riccati equation endogenously arises as in our model.

[8]Equivalently,

$$p(t) = \frac{p^* + \frac{2A+Bp^*}{\sqrt{4AC-B^2}} \tan\left((t-t^*)\sqrt{4AC-B^2}/2\right)}{1 - \frac{2Cp^*+B}{\sqrt{4AC-B^2}} \tan\left((t-t^*)\sqrt{4AC-B^2}/2\right)}.$$

Notice that the solution is not well-defined if $B^2 = 4AC$. In that case, the solution is

$$p(t) = \frac{1}{\frac{2C}{2Cp^*+B} - C(t-t^*)} - \frac{B}{2C}$$

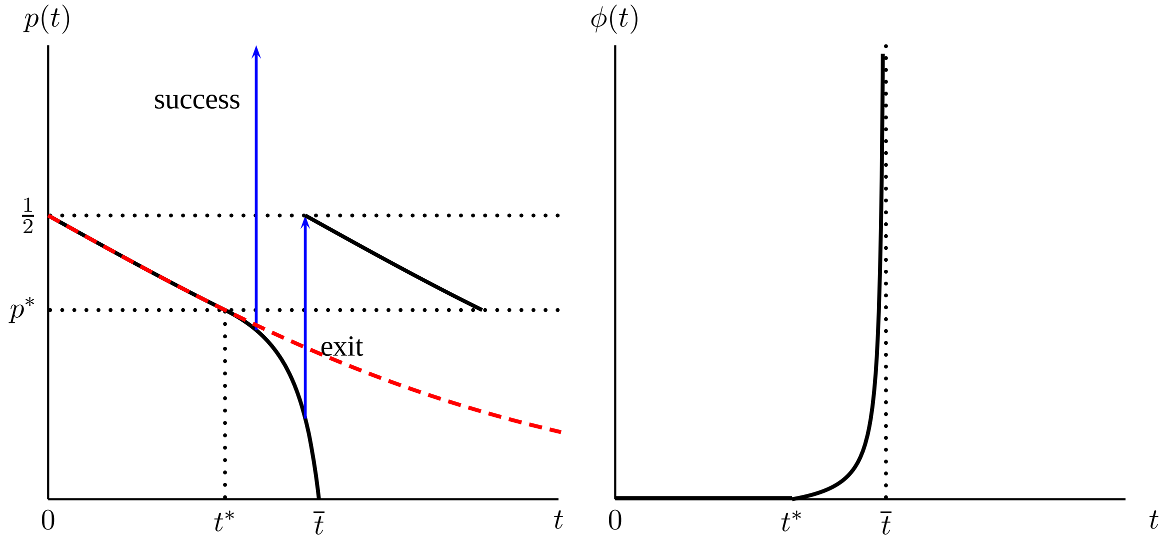

Figure 2: The players' beliefs $p(t)$ (left) and exit rate $\phi(t)$ (right) at time t, conditional on no success and no exit by the other player. The dashed line in the left panel represents the players' beliefs in the absence of social learning (i.e., $p_i e^{-\lambda t}/(p_i e^{-\lambda t} + 1 - p_i)$).

Figure 2 depicts the resulting equilibrium dynamics. As in the positive correlation case, the players, conditional on no success and no exit, exit at a positive rate from time t^*. Unlike in the positive correlation case, the players' beliefs $p(t)$ decrease, while their exit rate $\phi(t)$ increases, over time. As explained above, under negative correlation, both no success (private learning) and no exit by player j (social learning) make player i more pessimistic about his type, whereas under positive correlation, they work in the opposite direction and, in equilibrium, exactly cancel each other out. In order to compensate increasing pessimism, the players exit at an increasing rate (see the right panel of Figure 2). In other words, despite the fact that their beliefs fall below p^*, they are willing to stay in the game, because they expect to learn more from the other player's action. These two effects reinforce each other: as $p(t)$ decreases, $\phi(t)$ must increase. This pushes down $p(t)$ even further, which in turn leads to even larger $\phi(t)$. These effects grow exponentially fast and result in the players' beliefs $p(t)$ converging to 0 and the exit rate $\phi(t)$ converging to infinity in finite time (time $\bar{t}$). Notice that this does not mean that the players' beliefs would indeed become equal to 0. It simply means that one player exits with probability 1 before time $\bar{t}$.

5 Asymmetric Negative Correlation

We now consider the asymmetric case where the players assign different prior probabilities to the event that their type is good. Without loss of generality, we assume that player 1 is more likely to be good than player 2, that is, $p_1 > p_2 (= 1 - p_1)$. In order to avoid triviality, we also assume that

$p_2 > p^*$: otherwise, player 2 exits immediately.

For each $i = 1, 2$, we denote by $F_i(t)$ the probability that player i exits by time t, conditional on no success and no exit by player j. We also let $\phi_i(t)$ denote, whenever possible, player i's exit rate at time t (i.e., $\phi_i(t) = dF_i(t)/(1-F_i(t))$) and $p_i(t)$ denote the probability that player i assigns to the event that his type is good at time t, conditional on no success and no exit by player j.

As in the symmetric case, in equilibrium the two distribution functions F_1 and F_2 have a common convex support: otherwise, each player's best response is a pure strategy. In addition, the minimum of the support must be equal to $t^* = t_2^*(< t_1^*)$: it is the point at which player 2's belief $p_2(t)$ reaches p^*, and thus social learning must occur for him not to exit immediately. Unlike in the symmetric case, player 1's belief $p_1(t^*)$ exceeds p^*. This raises a subtle issue. Player 1 strictly prefers staying to exiting whenever $p_1(t) > p^*$, while player 1's exit rate must be positive (i.e., player 1's action must be informative) in order for player 2 to be willing to stay whenever $p_2(t) \leq p^*$. As shown shortly, this issue can be resolved by player 2's exiting with a positive probability at time t^*, as it makes player 1's belief $p_1(t)$ drop below p^* instantly, conditional on no exit by player 2.

We first solve for player 2's conditional belief $p_2(t)$ and player 1's exit strategy $F_1(t)$ (equivalently, $\phi_1(t)$). Since player 1 never exits with a positive probability (i.e, the distribution function F_1 has no atom anywhere), these two can be derived just as in the symmetric case. We then characterize player 1's conditional belief $p_1(t)$ and player 2's exit strategy $F_2(t)$.

Player 2's equilibrium belief and player 1's equilibrium exit strategy. Given player 1's exit strategy $\phi_1(t)$, conditional on no success and no exit by player 1, player 2's belief evolves as follows:

$$p_2(t) = \frac{p_2 e^{-\lambda t} e^{-\int_0^t \phi_1(y)dy}}{p_2 e^{-\lambda t} e^{-\int_0^t \phi_1(y)dy} + (1-p_2)\int_0^\infty e^{-\int_0^{\min\{x,t\}} \phi_1(y)dy} d(1-e^{-\lambda x})}.$$

Arranging the terms as in the symmetric case, it follows

$$\dot{p}_2(t) = -\lambda p_2(t)(1-p_2(t)) - \phi_1(t)p_2(t)\frac{p_2 - p_2(t)}{p_2}. \quad (7)$$

Observe that equation (3) is a special case of this equation, with $p_2 = 1/2$.

As in the symmetric case, player 2 must remain indifferent between staying and exiting whenever $t \in (t_2^*, \bar{t})$. Therefore,

$$c = \frac{p_2(t)}{p_2}\phi_1(t)V(p_2) + p_2(t)\lambda(v + V(1)). \quad (8)$$

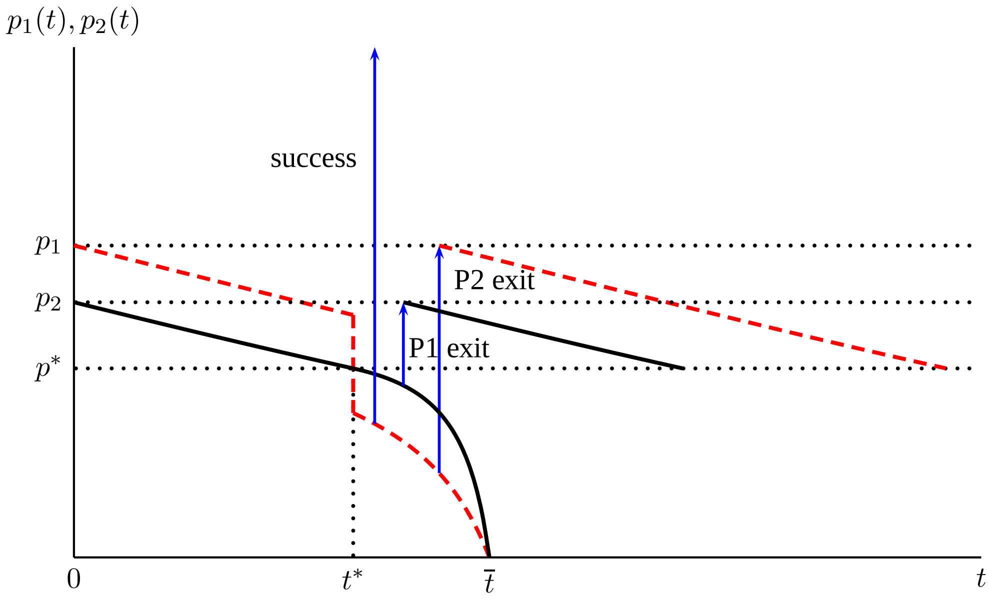

Figure 3: The players' beliefs conditional on no success in the asymmetric case. The dashed line is for player 1, while the solid line is for player 2.

This equation corresponds to equation (4) in the symmetric case. The only differences are that the probability that player 2 assigns to the event that player 1 has not succeeded by time t is equal to $p_2(t)/p_2$, instead of $2p_2(t)$, and that once player 1 exits (and thus player 2 knows that player 1 has not succeeded as well), player 2 updates his belief to p_2, instead of $1/2$.

Combining equations (7) and (8) leads to the following Riccati equation:

$$\dot{p}_2(t) = -\lambda p_2(t)(1 - p_2(t)) - \frac{c - p_2(t)\lambda(v + V(1))}{V(p_2)}(p_2 - p_2(t)). \tag{9}$$

As in the symmetric case, a closed-form solution can be obtained by applying Lemma 1, together with the boundary condition $p_2(t^*) = p^*$. Given the solution $p_2(t)$, player 1's equilibrium exit strategy $\phi_1(t)$ can also be explicitly derived. As in the symmetric case, $p_2(t)$ converges to 0 in finite time.

Player 1's equilibrium belief and player 2's equilibrium exit strategy. Following the same steps as above, we obtain the following equations: whenever $t \in (t^*, \bar{t})$,

$$\dot{p}_1(t) = -\lambda p_1(t)(1 - p_1(t)) - \phi_2(t)p_1(t)\frac{p_1 - p_1(t)}{p_1}, \tag{10}$$

and

$$c = \frac{p_1(t)}{p_1}\phi_2(t)V(p_1) + p_1(t)\lambda(v + V(1)). \tag{11}$$

Therefore, we again obtain a Riccati equation for $p_1(t)$:

$$\dot{p}_1(t) = -\lambda p_1(t)(1 - p_1(t)) - \frac{c - p_1(t)\lambda(v + V(1))}{V(p_1)}(p_1 - p_1(t)). \tag{12}$$

The difference from equation (9) lies in the boundary condition, because, as explained above, $\lim_{t\to t^*-} p_1(t) \neq p^*$. A necessary condition comes from the initial observation that the distribution functions F_1 and F_2 must have a common support. In particular, denote by $\bar{t}$ the upper bound of the support. Then, it must be that $p_1(\bar{t}) = p_2(\bar{t}) = 0$. Intuitively, if player j, conditional on no success, exits with probability 1 by time $\bar{t}$, then player i has no reason to stay beyond $\bar{t}$ and must assign probability 1 to his type being bad (player j's type being good), conditional on no success.

The exact value of $\bar{t}$ can be calculated from the solution to equation (9) (i.e., the value such that $p_2(\bar{t}) = 0$). Then, the condition that $p_1(\bar{t}) = 0$ can be applied to explicitly solve equation (12). Given the solution $p_1(t)$ over the interval $(t^*, \bar{t})$, player 2's equilibrium exit rate $\phi_2(t)$ can be obtained from equation (11). A necessary condition for $p_1(t)$ is that $p_1(t) < p^*$ for any $t \in (t_2^*, \bar{t}]$: otherwise, player 1 strictly prefers staying. This condition is guaranteed because $p_1(\bar{t}) = p_2(\bar{t}) = 0$, while $p_2(t)$ decreases faster than $p_1(t)$ before t^*: if $p_1(t) = p_2(t)$, then $|\dot{p}_1(t)| < |\dot{p}_2(t)|$, because $V(p_1) > V(p_2)$. Therefore, $p_1(t)$ must always stay below $p_2(t)$. Intuitively, player 1, due to the difference in prior beliefs, obtains more from player 2's exit than player 2 does from player 1's exit. Therefore, for them to be simultaneously indifferent between staying and exiting, player 1 must remain more pessimistic than player 2.

It remains to pin down player 2's exit probability at time t^*. Let $p_1^-(t^*) \equiv \lim_{t\to t^*-} p_1(t) (> p^*)$. Player 1's belief must jump down from $p_1^-(t^*)$ to $p_1(t^*)$. Conditional on player 1 being good, player 2 never succeeds and, therefore, exits with probability $F_2(t^*)$. Conditional on player 2 being good, player 2 exits only when he does not succeed by time t^* and, therefore, with probability $e^{-\lambda t^*} F_2(t^*)$. It follows that the probability that player 2 exits at time t^*, $F_2(t^*)$, must satisfy

$$p_1(t^*) = \frac{p_1^-(t^*)(1 - F_2(t^*))}{p_1^-(t^*)(1 - F_2(t^*)) + (1 - p_1^-(t^*))(1 - e^{-\lambda t^*} F_2(t^*))}. \tag{13}$$

We summarize all the results in the following proposition.

Proposition 4 *In the model with negative correlation (and $p_1 \geq p_2$), there exits a unique equilibrium. No player exits until time $t^* = t_2^*$. At time t^*, player 2 exits with probability $F_2(t^*)$ (as calculated in equation* (13)*), which lowers player 1's belief $p_1(t)$ below p^*. After time t^*, the players' beliefs, conditional on no success and no exit, evolve as in equations* (9) *and* (12)*, and each* *player exits at a positive rate as given in equations* (8) *and* (11). *Player 2's expected payoff is identical to that in Proposition 1, while player 1 obtains a strictly higher expected payoff than in Proposition 1 whenever* $p_1 > p_2 > p^*$.

Proof. See the appendix. ■

The result shows that social learning can be valuable under asymmetric negative correlation. More importantly, it illustrates when, and to whom, social learning is beneficial. As shown in the positive correlation case and the symmetric case, social learning does not improve the players' welfare if it occurs gradually. Intuitively, gradual social learning causes excessive delay, which offsets the benefit of social learning. If social learning occurs fast (as at time t^* in the asymmetric case), then it allows a player to enjoy the benefit, without incurring any delay cost. The player who enjoys the benefit is the one who is more willing to stay in the game and, therefore, relies less on social learning.

6 Discussion

In this section, we explain how to extend our analysis in various dimensions. For simplicity, we restrict attention to the symmetric case where the players begin with an identical probability of being good.

6.1 Imperfect Negative Correlation

We first consider the case of imperfect negative correlation, in which both players may be bad. In other words, now there are three possibilities: (i) player 1 is good, while player 2 is bad. (ii) player 1 is bad, while player 2 is good. (iii) both players are bad.[9] We denote by p_0 the prior probability of the last event. Since we consider the symmetric case, this means that the prior probability that each player is good (and the other is bad) is given by $p_i = (1 - p_0)/2$.

The equilibrium structure is similar to that of perfect negative correlation in Section 4. Let t^* be the point at which the players' beliefs reach p^* in the absence of social learning (that is, $p^* = p_i e^{-\lambda t^*}/(p_i e^{-\lambda t^*} + 1 - p_i)$ for $i = 1, 2$). Then, the players do not exit until time t^*, while they exit at a positive rate over the interval $[t^*, \bar{t})$.

There are two differences regarding buyers' conditional beliefs. First, player i's belief, conditional on no success and no exit by player j, decreases faster than under perfect negative correlation. This is because player i's own failure indicates not only the possibility that player j is good,

[9]One can consider the opposite case where both players may be good. The analysis is almost identical to the one here and, therefore, omitted.

but also the possibility that both players are bad. Formally, if $t \in (t^*, \bar{t})$, then player i's conditional belief evolves according to

$$p(t) = \frac{1}{2 + \int_0^t e^{\int_x^t (\lambda + \phi(y))dy} \lambda dx + \frac{2p_0}{1-p_0} e^{\lambda t}},$$

which implies that

$$\dot{p}(t) = -\lambda p(t)(1 - p(t)) - \phi(t)p(t) \left(1 - 2p(t) \left(1 + \frac{p_0}{1 - p_0} e^{\lambda t}\right)\right). \tag{14}$$

Notice that equation (14) coincides with equation (3) when $p_0 = 0$, and the right-hand side is decreasing in p_0.

Second, player i's belief following player j's exit depends on the time player j exits: recall that player i's belief always goes back to p_i after player j's exit under perfect negative correlation. This is because the probability that both players are bad, conditional on no player's success, is strictly increasing over time. Formally, if player j exits at time t, then player i's conditional belief updates to

$$\frac{p_i e^{-\lambda t}}{p_0 + p_i e^{-\lambda t} + p_j e^{-\lambda t}} = \frac{1}{\frac{2p_0}{1-p_0} e^{\lambda t} + 2} < \frac{1}{2}.$$

Combining this with the fact that player i must remain indifferent between staying and exiting whenever $t \in (t^*, \bar{t})$,

$$c = \left(2 + \frac{2p_0}{1 - p_0} e^{\lambda t}\right) p(t)\phi(t)V \left(\frac{1}{2 + \frac{2p_0}{1-p_0} e^{\lambda t}}\right) + p(t)\lambda(v + V(1)). \tag{15}$$

Notice that if $p_0 = 0$, then this equation reduces to equation (4).

As in Section 4, combining equations (14) and (15) yields a Riccati equation for $p(t)$. By standard arguments, there exists a unique solution to the equation. However, we are not able to derive an explicit solution. Technically, this is because, although the equation is similar to equation (5), its coefficients are time-varying, in which case closed-form solutions are known only for a limited class of Riccati equations.

6.2 More Players

We have restricted attention to the case where there are only two players. As explained below with the case of three players, the analysis becomes significantly more complicated in the general N-player case. This is a severe limitation of our analysis, given that Murto and Välimäki (2011) provide a characterization for the general case. A major difference is that buyers' conditional beliefs in the second phase remain constant even in the general case under positive correlation, which simplifies the analysis significantly as demonstrated in Section 3.2, but vary over time in an intricate way under negative correlation.

To be concrete, consider the case of three players in which only one of them is good. Let t^* be the point at which the players' conditional beliefs become equal to p^* in the absence of social learning (i.e., t^* is the value such that $p^* = e^{-\lambda t^*}/(2 + e^{-\lambda t^*})$). Then, no player exits until time t^*. After time t^* until one player exits, they exit at a positive rate. Denote by $\phi(t)$ the symmetric exit rate of the players. Then, their conditional beliefs evolve according to

$$p(t) = \frac{1}{3 + 2\int_0^t e^{\int_x^t (\lambda+\phi(y))dy} \lambda dx}.$$

Suppose one player exits at time t. This reveals that the player has not succeeded yet. Then, the remaining players update their conditional beliefs to

$$p(t) = \frac{1}{3 + \int_0^t e^{\int_x^t (\lambda+\phi(y))dy} \lambda dx}.$$

Once a player exits, the game turns to a two-player game. Importantly, this subgame involves imperfect negative correlation, because there is a positive probability that the player who exited is actually good. As explained above, this problem is technically a lot more challenging than our baseline model, which also points to the difficulty of further characterizing the three-player case.

Nevertheless, it is possible to infer the limiting equilibrium outcome as the number of players tends to infinity. Suppose there are N players and it is common knowledge that $M(< N)$ players are good. Now let both M and N tend to infinity, while keeping its ratio M/N constant. In the limit, the problem becomes trivial, because, by the law of large numbers, correlation among players' types disappears, and thus there is no social learning. In other words, each player's problem reduces to the single-player problem in Section 3.1. The problem becomes non-trivial if there is aggregate uncertainty about the ratio M/N, as in Murto and Välimäki (2011). However, again by the law of large numbers, the problem becomes identical to that of Murto and Välimäki (2011) in the limit as N tends to infinity.

6.3 Re-entry

Negative correlation gives rise to an incentive for the player who exits first to re-enter the game later. To be specific, consider the symmetric case in Section 4 and suppose player j exited first. As shown in Section 4, there is a positive probability that player i exits later. At that point, player j's belief is equal to $1 - p^*$, which is even above $1/2$: although it is common knowledge that both players have not succeeded, player i has experimented longer than player j, and thus player j is more likely to be good than player i. Therefore, player j is willing to re-enter the game as long as the re-entry cost does not exceed $V(1-p^*)$. Full characterization is fairly involved, mainly because the exiting player's expected payoff is now strictly positive (i.e., Lemma 1 in Murto and Välimäki (2011) no loner applies). However, the following result is straightforward to establish. Suppose the players can repeatedly re-enter the game if they pay a fixed cost $e > 0$ each time. If $e \geq V(1-p^*)$, then the equilibrium without re-entry in Proposition 3 remains as an equilibrium. If $e < V(1-p^*)$, then the players alternately re-enter the game until one player eventually succeeds and, therefore, stays forever.

Appendix

Proof of Proposition 2. Given player j's strategy $\phi_j(t)$, player i's belief, conditional on no success and no exit by player j, evolves according to

$$\begin{aligned} p_i(t) &= \frac{p_i e^{-\lambda t}\int_0^\infty e^{-\int_0^{\min\{x,t\}}\phi_j(y)dy}d(1-e^{-\lambda x})}{p_i e^{-\lambda t}\int_0^\infty e^{-\int_0^{\min\{x,t\}}\phi_j(y)dy}d(1-e^{-\lambda x}) + (1-p_i)e^{-\int_0^t \phi_j(y)dy}} \\ &= \frac{\int_0^t e^{\int_x^t(\lambda+\phi_j(y))dy}\lambda dx + 1}{\int_0^t e^{\int_x^t(\lambda+\phi_j(y))dy}\lambda dx + 1 + \frac{1-p_i}{p_i}e^{2\lambda t}}. \end{aligned}$$

Differentiating with respect to t and arranging the terms,

$$\dot{p}_i(t) = -p_i(t)(1-p_i(t))\lambda + \phi_j(t)(1-p_i(t))\frac{((1-p_i)e^{2\lambda t}+p_i)p_i(t)-p_i}{(1-p_i)e^{2\lambda t}}.$$

Applying Lemma 3 in Murto and Välimäki (2011), $p_i(t)$ should stay constant once it reaches p^*. Therefore, it follows that

$$\phi_1(t) = \phi_2(t) = \frac{\lambda p^*(1-p_i)e^{2\lambda t}}{((1-p_i)e^{2\lambda t}+p_i)p^*-p_i}.$$

The result that if player j exits, then player i immediately follows comes from the fact that player i's belief, conditional on no success, jumps down to

$$p_i(t) = \frac{p_i e^{-2\lambda t}}{p_i e^{-2\lambda t}+1-p_i} \leq \frac{p_i e^{-2\lambda t^*}}{p_i e^{-2\lambda t^*}+1-p_i} < \frac{p_i e^{-\lambda t^*}}{p_i e^{-\lambda t^*}+1-p_i} = p^*.$$

■

Proof of Propositions 3 and 4.

(i) No player exits until time $t^* \equiv \min\{t_1^*, t_2^*\}$.

Define $\tilde{t}_i \equiv \inf\{t : p(t) \le p^*\}$, and $\tilde{t} \equiv \min\{\tilde{t}_1, \tilde{t}_2\}$. Since no player exits until $\tilde{t}$, $p_i(t) = p_i e^{-\lambda t}/(p_i e^{-\lambda t} + 1 - p_i)$ for any $t < \tilde{t}$. It is then immediate that $\tilde{t} = t^*$.

(ii) If $t > t^*$, then $p_i(t) \le p^*$ for both $i = 1, 2$.

Suppose $p_i(t) > p^*$ for some $t > t^*$. Since $p_i(t)$ is always decreasing over time and player i never exits when $p_i(t) > p^*$ (Lemma 1 in Murto and Välimäki (2011)), this means that player j does not learn from player i's behavior over the interval $[t^*, t]$. Given this, player j, conditional on no success, prefers exiting immediately at time t^*: formally, $p_j(t') < p^*$ whenever $t' \in (t^*, t)$, and thus $c = p^*\lambda(v+V(1)) > p_j(t)\lambda(v+V(1))$ (see equation (4)). But this implies that $p_i(t^*+dt) = 0$ (note that $p_i(t)$ is player i's belief conditional on no success and no exit by player j), which is a contradiction.

(iii) The two distribution functions F_1 and F_2 have a common support of the form $[t^*, \bar{t}]$ for some $\bar{t}(> t^*)$ and are continuous. Finally, $F_1(t^*)F_2(t^*) = 0$.

Let $\underline{t}_i$ and $\bar{t}_i$ denote the lower bound and the upper bound of the support of F_i. Applying the same reasoning as in (ii), $\underline{t}_1 = \underline{t}_2 = t^*$. Now suppose $\bar{t}_i > \bar{t}_j$. In this case, player i does not learn from player j's behavior after $\bar{t}_j$. Since $p_i(t) \le p^*$, conditional on no success, he exits immediately, which is a contradiction.

Now we show that the common support of F_1 and F_2 is the interval $[t^*, \bar{t}]$. Suppose $F_i(t)$ is constant on $[t^1, t^2) \subset [t^*, \bar{t})$. In this case, by the same reasoning as in (ii), player j, conditional on no success, exits immediately at t_1, which is a contradiction. Now suppose $F_i(t)$ has an atom at $t \in (t^*, \bar{t})$. In this case, player j has no incentive to exit close to t, that is, there exits $\varepsilon > 0$ such that $F_j(t)$ is constant on $[t-\varepsilon, t)$, which is a contradiction.

The no-atom result above does not apply to t^*. However, if $F_i(t^*) > 0$, then player j clearly strictly prefers waiting an instant more than exiting immediately, and thus $F_j(t^*) = 0$.

(iv) All the results here imply that an equilibrium necessarily takes the structure employed in the main text. The equilibrium uniqueness then follows follow from an explicit equilibrium construction in the main text.

(v) Denote by $V_1(t)$ player 1's expected payoff at time t. Whenever $p_1 > p_2$ (equivalently, whenever $F_2(t^*) > 0$), player 1's expected payoff $V_1(0)$ exceeds $V(p_1)$.

After time t^*, since exit is always an optimal strategy, player 1's expected payoff remains equal to 0. Therefore,

$$V_1(t^*) = \left(p_1^-(t^*) + (1 - p_1^-(t^*))e^{-\lambda t^*}\right) F_2(t^*)V(p_1).$$

Since the value function $V(\cdot)$ is convex, $V_1(t^*) > V(p_1^-(t^*))$. The desired result then follows from the fact that both $V_1(t)$ and $V(t)$ decrease according to the same law of motion over the interval $t \in [0, t^*)$ ($rV_1(t) = -c + \lambda p_1(t)(v + V(1) - V_1(t)) + \dot{V}_1(t)$ and $rV(t) = -c + \lambda p_1(t)(v + V(1) -$

$V(t)) + \dot{V}(t)$ with $p(t) = p_1 e^{-\lambda t}/(p_1 e^{-\lambda t} + 1 - p_1))$ and, therefore, cannot cross each other.

■

References

Bolton, Patrick and Christopher Harris, "Strategic experimentation," *Econometrica*, 1999, *67* (2), 349–374.

Bonatti, Alessandro and Johannes Hörner, "Collaborating," *American Economic Review*, 2011, *101* (2), 632–663.

Keller, Godfrey and Sven Rady, "Strategic experimentation with Poisson bandits," *Theoretical Economics*, 2010, *5* (2), 275–311.

_ , _ , and Martin Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, 2005, *73* (1), 39–68.

Klein, Nicolas and Sven Rady, "Negatively correlated bandits," *Review of Economic Studies*, 2011, *78* (2), 693–732.

Ljungqvist, Lars and Thomas J. Sargent, *Recursive macroeconomic theory*, MIT press, 2004.

Murto, Pauli and Juuso Välimäki, "Learning and information aggregation in an exit game," *Review of Economic Studies*, 2011, *78* (4), 1426–1461.

Nawalkha, Sanjay K., Gloria M. Soto, and Natalia A. Beliaeva, *Dynamic Term Structure Modeling: The Fixed Income Valuation Course*, John Wiley & Sons, 2007.

Rosenberg, Dinah, Eilon Solan, and Nicolas Vieille, "Social learning in one-arm bandit problems," *Econometrica*, 2007, *75* (6), 1591–1611.