Strategic Experimentation On A Common Threshold

Yi Chen*

Dec, 2015

Abstract

A dynamic game of experimentation is examined where players search for an unknown threshold. Players contribute to the rate of decline in a state variable, and the game ends with a costly breakdown once the state falls below the threshold. In the unique symmetric pure-strategy stationary Markov equilibrium, the state decreases gradually over time and settles at a cutoff level asymptotically, conditional on no breakdown. The cutoff depends on the patience, the cost of the breakdown, and the prior distribution of the threshold, but not on the number of players. In a discrete-time version of the game, the equilibrium time path of the state converges to that of the continuous-time model when period length tends to zero.

Keywords: Experimentation; Free-Riding; Markov Perfect Equilibrium; Information Externality; Spatial Learning

JEL Codes: D81, D82, D83

1 Introduction

Learning through experimentation is a common practice in many aspects of the economy. In the literature, a large number of models used the "bandit" framework that agents learn about the quality of one or several "arms" through a potential stochastic process (Cripps, Keller, and Rady (2005), Bolton and Harris (1999), Bonatti and Hörner (2011), etc.). However, there are other situations in the economy where people need to learn about an unknown threshold among arms rather than the prospect of a given arm. Whenever the outcome as a function of the arm being operated takes a discontinuous all-or-none form, the threshold that triggers the discrete change in outcome is an important

*Yale University. Email: y.chen@yale.edu.

parameter to learn. Situations of this nature are common in our economy. Manufacturers try to pass some vague regulatory criteria when setting the quality of their products. Countries with similar environmental conditions face a threshold level of effort to control pollution, below which a costly ecological disaster happens. Risk management sector of a bank struggle to search for the threshold level of risk exposure above which a financial distress occurs, but low risk is costly in terms of profit.

In order to capture collective learning situations of this sort, I present a model that abstracts the key feature of the "binary outcome" out of reality: as the input (the arm being operated) continuously decreases, the outcome (the payoff from that arm) has a downward jump at some unknown threshold. Using terminology from the bandit literature, my model features a continuum of arms, although there are two important differences from that literature. First, the arms in my model are dependent in that if one arm is bad (below the threshold), then all arms below it are also bad. Second, the continuum of arms in my model poses a "spatial" dimension that is the main object of learning. In a typical exponential bandit problem, learning is slow in the sense that it takes time for belief to evolve regarding a specific arm. To some extent, this way of learning is "chronological", learning is slow and the key decision is *when* to switch arms. In my model, in contrast, learning on a specific arm is assumed to be instant but learning for the threshold is still slow because of the continuous action space. I call it "spatial" learning since the key decision parameter is *how fast* to skim down the spectrum of arms.

Specifically, I first consider a continuous-time multi-player team problem. There is a common starting "effort level". At each time, all players receive a fixed flow benefit, and chooses a "contribution" to the decline of effort. The higher the common effort, the larger the flow cost. The process continues until the current common effort falls below a threshold level. The threshold is fixed from the beginning, but is unknown to the players. At that moment, the game ends, the team is hit with a lumpy cost, and then receives a terminal payoff that is equal to the present discounted value of staying slightly above the threshold thereafter. As is solved in Section 3, the solution to the team problem entails a very fast decline in common effort as long as it is above certain cutoff, after which it stays constant.

In contrast, a multi-player game gives qualitatively different dynamics. Whenever the threshold is hit, the lumpy cost is shared among players proportional to the relative contribution at that moment. There are information externalities among players: the history of play is common knowledge, and thresholds they face are the same. In other words, contribution from a player benefits others by providing more learning about the threshold, but larger contribution means higher share in the lumpy cost, which is not internalized. As a result, players tend to contribute less to the decline of effort, free riding on the information provided by others. The game is shown to have a unique symmetric Markov equilibrium, in which the common effort declines smoothly over time–faster in the beginning and then slows down towards a long-run asymptotic level. Interestingly, this asymptotic level is the same as the cutoff level in the team problem.

The paper connects to several branches of existing literature. First, it belongs to the tradition of experimentation a la Rothschild (1974). Later, exponential bandit (Cripps, Keller, and Rady (2005), Bonatti and Hörner (2011)) and Brownian bandit (Bolton and Harris (1999), Moscarini (2005)) were most extensively studied. This paper contributes to the literature with a less studied form of learning, where the dynamic choice is among the continuum of arms available. Chronologically learning may never stop, but spatially it should end outside some "inaction region".

Second, the paper contributes to the studies of dynamic games. Admati and Perry (1991) and Matthews (2013) studied the dynamic problem of contribution to a public good, without uncertainty and learning. Maskin and Tirole (1988a,b) investigated a dynamic stationary duopoly game. Lancaster (1973), Levhari and Mirman (1980) analyzed the dynamic game of the extraction of a common resource. My model fits well into this category of games, although in a continuous-time form. Section 5 explores the property of a discrete-time version of the model, and establishes convergence to the main model.

Moreover, the paper is closely related to three papers on spatial learning. Bonatti and Hörner (2013) considered a single player learning problem with respect to an unknown threshold, where action space is finite and the consequence of taking a too low action is a Poisson process. Callander (2011) featured a spatial learning in which the realization of each trial is a sample point on a path of Brownian motion. Rob (1991) investigated an industrial organization model where a continuum of firms collectively learn the position of a kink on the demand curve by decisions of entry and exit. My work models a different underlying economic situation, and investigates the strategic interaction between several players learning the same object.

The rest of the paper is organized as follows. Section 2 lays out the settings of the main model. Before giving the analysis of the game, Section 3 considers the cooperative problem first, serving as a benchmark. Section 4 solves the strategic problem for the multi-player game. Section 5 describes the discrete time counterpart of the game, and shows the convergence of the solutions to that of the continuous time game. Section 6 extends the game in three directions. Section 7 concludes.

2 Model Setup

The Game Time is continuous. The duration of the game is random, and the realized duration is denoted by $\bar{t}$. There are $I \geqslant 1$ players labeled $i = 1, \ldots, I$. At each time $t \in \mathbb{R}_+$, Player i ($i = 1, \ldots, I$) chooses an action $v_i(t) \in [0, \overline{v}]$ (where $\overline{v} > 0$) if the game has not ended, and I write $v_i = \{v_i(t)\}_{t=0}^{\bar{t}}$ as the resulting path.

There is a state variable $x \geqslant 0$ which I interpret as the level of (collective) effort. The law of motion of x is as follows:

$$x(t) = x_0 - \int_0^t V(s)ds, \quad \text{where } V(t) \equiv \sum_{i=1}^{I} v_i(t) \tag{1}$$

for some initial value $x_0 > 0$. Hence, each individual $v_i(t)$ serves as an additive contribution to the total "rate of decline" in x.

The game features incomplete information. Specifically, there is a random state of the world $c \in [0, x_0)$. It is distributed with c.d.f. $F(\cdot)$ and a continuous density $f(\cdot)$ on support $[0, x_0]$. The realized value of c is fixed throughout, but unknown at the start. The game ends when $x(t) \leqslant c$ for the first time, if ever. In this sense, c is regarded as a *threshold*.[1] The duration of the game is hence a stopping time

$$\bar{t} = \inf\{t : x(t) \leqslant c\} \in \mathbb{R}_+ \cup \{+\infty\}.$$

Because $V(t)$ is bounded for any $t \in [0, \bar{t}]$, the time path of effort $x(\cdot)$ is continuous. Therefore, if $\bar{t} < \infty$, the threshold c is revealed to be $x(\bar{t})$ at the end of the game.

From now on I maintain the following assumptions on the distribution of c, unless explicitly mentioned otherwise.

Assumption 1 *(Monotone Hazard Rate)*
The inverse hazard rate $\frac{F(\cdot)}{f(\cdot)}$ is strictly increasing.

Assumption 2 *(Strongly Positive Density)*
The density $f(\cdot)$ is uniformly bounded away from zero, i.e. $\exists \underline{f} > 0$ s.t. $f(c) \geqslant \underline{f}$ for all $c \in [0, x_0]$.

Assumption 3 *(Lipschitz Continuous Density)*
The density $f(\cdot)$ is Lipschitz continuous, i.e. $\exists \kappa > 0$ s.t. $|f(x) - f(y)| \leqslant \kappa |x - y|$ for all $x, y \in$

[1] In the language of bandits, c is the threshold that divides the continuum of arms into "good" arms ($x > c$) and "bad" arms ($x \leqslant c$)

$[0, x_0]$.

Assumption 1 is a standard monotonicity assumption. Assumptions 2 and 3 rule out pathological solutions of the game.

Payoff comes in two forms: flow and lump. Flow cost consists of cost and benefit. Higher flow cost results from higher effort. With little loss of generality, the flow cost is simply x for everyone when the effort level is x.[2] Also, each player enjoys a fixed flow benefit $p > x_0$ at all times.

Lumpy payoffs come at the end of the game, if it ends in finite time. Given a fixed path $x(\cdot)$, the game ends at $\bar{t}$ ($\bar{t} = \infty$ means no breakdown) along with a costly *breakdown*. Let r be the common discount rate of the players. First, the breakdown brings lumpy cost $L > 0$, if ever, at time $\bar{t}$. It is assumed that $Lr < \frac{F(x_0)}{f(x_0)}$, i.e. the lumpy cost is relatively small. [3] The lumpy cost is divided among players as follows: Player i suffers $L\frac{v_i(\bar{t})}{V(\bar{t})}$, so that the share of cost equals the share of actions at that moment.[4,5] Second, each player receives a terminal benefit $\frac{p-c}{r} > 0$.[6]

In sum, if the breakdown occurs at time $t = \bar{t}$ which reveals threshold $c = x(\bar{t})$, then Player i's realized total present discounted payoff is

$$\underbrace{\int_0^{\bar{t}} (p - x(t))\, e^{-rt} dt}_{\text{flow before } \bar{t}} + \underbrace{\frac{p-c}{r} e^{-r\bar{t}}}_{\text{terminal benefit at } \bar{t}} - \underbrace{L\frac{v_i(\bar{t})}{V(\bar{t})} e^{-r\bar{t}}}_{\text{terminal cost at } \bar{t}},$$

where $x(t)$ is determined by (1).

[2]Actually, as long as the flow cost function $\phi(x)$ is Lipschitz continuous and strictly increasing in x, we can always redefine the state variable (the effort level) to be $\phi(x)$ instead of x, so that the cost is again equal to the state.

[3]As will become evident later, with this assumption at least some learning is worthwhile at the beginning.

[4]It is well-defined because the breakdown arrives only when $V > 0$.

[5]It deserves some explanation why the probability of suffering the lumpy cost is shared in this way. It is not some designed "cost sharing rule" that players agree upon; it is simply a natural implication if we view the continuous time game as a limit of discrete time games with very short period length. In the latter, players take turns to bring x downward from the previous level. If the turns shift frequently enough, then the probability of hitting the threshold during one's turn is approximately proportional to the change of x in her turn. The reader is referred to Section 5 for the description of a discrete time model.

[6]This form of terminal payoff is not crucial for results. Actually we can allow for fairly general functions, but the one used here has the following economic interpretation: the arrival of breakdown at $\bar{t}$ fully reveals the location of the threshold, so if there was a continuation game after the stopping time, the supremum of payoff for any player is a constant flow of $p-c$. The terminal benefit $\frac{p-c}{r} > 0$ is the present discounted value of this flow. It is positive because by assumption $p > x_0 \geqslant c$. One can also view this as a normalization because in this way the real cost of a breakdown is indeed L, instead of L plus some change of continuation value.

Information and Beliefs I assume perfect monitoring of players' actions. A *plausible* history of length t is denoted by $h^t \equiv \{v_i(s) : s < t\}_{i=1}^I$ such that $x(t) > 0$, and the set of all plausible histories of length t is $\mathcal{H}^t$. The space of all plausible histories is $\mathcal{H} \equiv \cup_{t \in \mathbb{R}_+} \mathcal{H}^t$. Beginning with the common prior $F(\cdot)$ at $t = 0$, players update their beliefs about the distribution of the threshold provided that no breakdown has occurred so far. Since for every x the outcome $\mathbb{1}_{\{x>c\}}$ is revealed immediately, updating process is a mere truncation at the top of the prior distribution.

Solution Concept A pure strategy of Player i is a measurable map from $\mathcal{H}$ into $[0, \overline{v}]$. I focus on stationary Markov perfect equilibrium in pure strategies so that the strategy profile depends only on the payoff-relevant state variable $x \in (0, x_0]$, but not on calendar time or on how the state x is reached. Formally, for all $i = 1, \dots, I$ a stationary Markov strategy of Player i is a measurable map $\nu_i : (0, x_0] \to [0, \overline{v}]$. A profile of Markov strategies is denoted by $\nu \equiv \{\nu_i\}_{i=1}^I$. The time path of x follows the ordinary differential equation below

$$\frac{dx}{dt} = -\sum_{i=1}^{I} \nu_i(x), \quad x(0) = x_0. \tag{2}$$

To ensure the existence and uniqueness of $x(\cdot)$, I restrict attention to strategies such that for all $i = 1, \dots, I$: (a) ν_i is left continuous for all $i = 1, \dots, I$, and (b) ν_i is piecewise Lipschitz continuous with at most finitely many jumps. With these restrictions on the strategy profile, the path of x as is uniquely determined by ν, so that the stopping time $\bar{t}$ is well-defined.

Given the strategy profile of other players, Player i aims to maximize the expect payoff

$$w_i(\nu) = \mathbb{E}\left[\int_0^{\bar{t}} (p - x(t))\, e^{-rt} dt + \frac{p-c}{r} e^{-r\bar{t}} - L \frac{\nu_i(c)}{\mathcal{V}(c)} e^{-r\bar{t}}\right] \tag{3}$$

where $x(t)$ is determined by ν via (2), $\bar{t}$ is pinned down by both ν and c, and the expectation is taken over c. A stationary Markov equilibrium in pure strategies (henceforth pure MPE) is a profile ν that constitutes a Nash equilibrium for any initial state $x \in (0, x_0]$.

3 Cooperative Problem

In this section I discuss the cooperative problem as a benchmark of the first best. The Hamilton-Jacobi-Bellman equation (henceforth HJB) of the social planner is given. Its analysis yields the solution to the problem, summarized in Proposition 1. Several implications from the result are discussed after that.

Denote $\overline{V} \equiv I\overline{v}$ as the upper limit on aggregate action, and denote $U_I(x)$ as the per-capita value function in the I-player cooperative problem. Formally, I solve the following optimization problem.

$$\begin{aligned} U_I(x_0) &= \frac{1}{I}\max_{\nu(\cdot)} \sum_{i=1}^{I} w_i(\nu) \\ &= \max_{\nu(\cdot)} \mathbb{E}\left[\int_0^{\bar{t}} (p - x(t))\, e^{-rt} dt + \frac{p-c}{r} e^{-r\bar{t}} - \frac{L}{I} e^{-r\bar{t}}\right] \end{aligned}$$

such that x evolves according to (2), $\bar{t}$ equals the stopping time determined by x, and $\nu_i(x) \in [0, \overline{v}]$ for all $x \in [0, x_0]$. Since x and $\bar{t}$ both depend only on $\mathcal{V}(\cdot) = \sum_{i=1}^{I} \nu_i(\cdot)$, it is sufficient for the social planner to control the aggregate strategy $\mathcal{V}(\cdot)$ only.

Recall that with the assumption on L, we have $\frac{F(x_0)}{f(x_0)} > Lr \geqslant \frac{Lr}{I}$. On the other hand, $\lim_{x\to 0} \frac{F(x)}{f(x)} = 0 < \frac{Lr}{I}$.[7] By Intermediate Value Theorem and Assumption 1, there exists a unique solution to the equation $\frac{F(x)}{f(x)} = \frac{Lr}{I}$. Let x_I^* denote the solution, where the subscript indicates the number of players in the problem.

For $x \in (0, x_0]$, if U_I is differentiable at x (as will be verified), then the HJB equation for the cooperative problem is

$$rU_I(x) = (p - x) + \max_{V \in [0, \overline{V}]} V \left\{ \frac{f(x)}{F(x)} \left(\frac{p-x}{r} - U_I(x) - \frac{L}{I} \right) - U_I'(x) \right\}. \tag{4}$$

The first term on the right-hand side is the flow payoff (benefit and cost) per capita. The second term is the maximum of a linear function of V, and the bracket multiplying V summarizes the benefit and cost of taking aggregate action V. The first term in the bracket is the expected loss upon hitting the threshold, equalling the hazard rate $f(x)/F(x)$ times

[7]The equality holds for the following reason. By Assumption 1 $\frac{F(x)}{f(x)}$ is increasing, and by definition it is bounded below by 0. Hence $\lim_{x\to 0} \frac{F(x)}{f(x)}$ exists and is non-negative. If $\lim_{x\to 0} \frac{F(x)}{f(x)} = a > 0$, then $f(x) \leqslant F(x)/a$ for all $x \geqslant 0$, and with initial condition $F(0) = 0$, we get $F(x) = 0$ for all $x \geqslant 0$, a contradiction. Hence, $\lim_{x\to 0} \frac{F(x)}{f(x)} = 0$.

the net loss per capita conditional on breakdown. The second term in the bracket captures the gains from learning by bringing x down to "$x - Vdt$", when there is no breakdown in the next moment.

The linearity of the HJB leads to a corner solution: $V = 0$ or $V = \overline{V}$. When $V = 0$, we have $U_I(x) = (p-x)/r$. When $V = \overline{V}$, (4) implies the following ordinary differential equation

$$\overline{V}U_I'(x) + \left(r + \frac{\overline{V}f(x)}{F(x)}\right)U_I(x) = p - x + \frac{\overline{V}f(x)}{F(x)}\left(\frac{p-x}{r} - \frac{L}{I}\right),$$

the solution to which is

$$U_I(x) = \frac{p-x}{r} + \frac{e^{-rx/\overline{V}}}{rF(x)}\left[C_1 + \int_0^x \left(F(s) - \frac{Lr}{I}f(s)\right)e^{rs/\overline{V}}ds\right], \tag{5}$$

for some $C_1 \in \mathbb{R}$. The right-hand side of the above consists of two parts. The first term is the payoff of staying at x forever (no learning), and the second term is the option value of learning. Value-matching and smooth-pasting pin down the constant C_1 and the cutoff effort level at which V switches from $\overline{V}$ to zero.

Proposition 1 *In the cooperative problem, the policy and the per-capita value function are*

$$V = \begin{cases} \overline{V} & \text{if } x \in (x_I^*, x_0] \\ 0 & \text{if } x \in (0, x_I^*] \end{cases},$$

$$U_I(x) = \begin{cases} \frac{p-x}{r} + \frac{1}{rF(x)}\int_{x_I^*}^x \left(F(s) - \frac{Lr}{I}f(s)\right)e^{-r(x-s)/\overline{V}}ds & \text{if } x \in (x_I^*, x_0] \\ \frac{p-x}{r} & \text{if } x \in (0, x_I^*] \end{cases}. \tag{6}$$

Proof. The value-matching condition requires $U_I(x_I^*) = (p - x_I^*)/r$, and the smooth-pasting boils down to $U_I'(x_I^*) = -1/r$. Plugging these two equations in (5) yields $\frac{F(x_I^*)}{f(x_I^*)} = \frac{Lr}{I}$. The optimality of the solution is guaranteed by Verification Theorem. ■

The solution (6) has several implications. First, there exists an inaction region (or "safety buffer") $[0, x_I^*]$. The fact that $x_I^* > 0$ means that the planner wants players to stop learning before they search through the entire spectrum of effort levels, even in the absence of a breakdown. Intuitively, there are cost and benefit when choosing the total speed of learning V. In (4), the benefit $-VU_I'(x)$ is the possible future gains from learning, and the cost $\frac{Vf(x)}{F(x)}\left(\frac{p-x}{r} - U_I(x) - \frac{L}{I}\right)$ is the risk of hitting the threshold at the next moment. As x becomes smaller, the hazard rate $f(x)/F(x)$ grows larger and eventually the cost outweighs the benefit. Intuitively, knowing that the range of the costly threshold is sufficiently narrowed down, the planner optimally calls off learning.

Second, first best allows for no procrastination in learning. In other words, the planner wants the players to reach x_I^* as fast as possible and then stop learning immediately once it is not worthwhile. Procrastination serves to push back the arrival of breakdown, but also necessitates a longer duration of high flow cost from effort. In the absence of breakdown, the time path of effort x is piecewise linear: skimming down from x_0 to x_I^* with maximum speed, followed by a constant level at x_I^* forever. As we will see in Section 4, procrastination naturally arises in non-cooperative situations.

Third, the reason that players switch from "risky" actions $V > 0$ to "safe" actions $V = 0$ is different from exponential bandits. In a two-armed exponential bandit with good news, players eventually switch from risky arm to safe arm because they are sufficiently pessimistic, in the sense that the current sacrifice in flow cost becomes too large compared to the potential benefit from learning. In the bad news case, on the other hand, once started learning is stopped only by a breakdown, and there is no voluntary switch on path. The spatial learning model considered here differs from both cases. It is in nature a bad news problem, and as the effort level decreases players grow more optimistic about the distribution of c (in the sense of first order stochastic dominance). However, unlike the exponential bandit with bad news, the learning stops when players get sufficiently optimistic. The seemingly paradoxical solution is explained by the fact that while the posterior distribution truncated at x first order stochastically dominates the one truncated at $x' < x$ so that globally speaking the state x' is better for the planner, the hazard rate of breakdown at x' is higher than the one at x (Assumption 1), which means that locally speaking the incentive to search further down is lower for state x'. Hence, it is the ordering of local incentives that determines the dynamics, not the global comparison of distributions. At some point, players stop learning when it becomes relatively too risky to search down, and at that point the current effort level is considered to be cheap enough.

One can also interpret the stopping property from the perspective of bandit problem with a continuum of correlated arms. Suppose a social planner faces a current state variable x, which is the highest arm whose output is uncertain. Heuristically, consider operating a "bunch of arms" in the interval $[x + dx, x]$ for the next instant of time dt. This bunch of arms, as a whole, is bad with probability $-dx\frac{f(x)}{F(x)}$. Hence, the experimentation generates a good outcome (no breakdown) with probability $1 + dx\frac{f(x)}{F(x)}$ saving future cost of $I\frac{-dx}{r}$ in total, and brings a bad outcome (breakdown) with probability $-dx\frac{f(x)}{F(x)}$ inflicting a damage of L. This marginal experimentation is profitable if and only if $\frac{F(x)}{f(x)} > \frac{Lr}{I}$. In summary, the logic for an individual arm is the same as exponential bandit, but learning

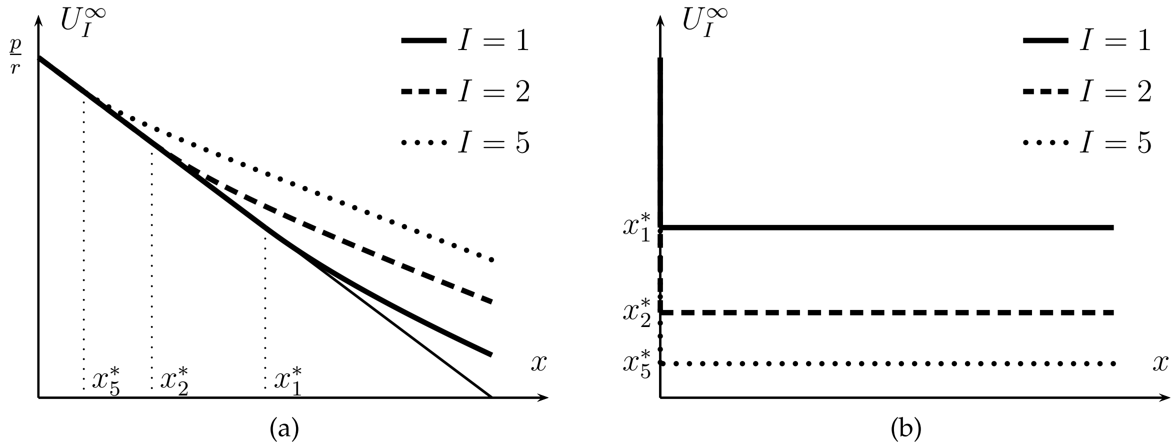

Figure 1: Cooperative Solution. (a) Value functions for 1, 2 or 5 players. (b) Time paths for 1, 2 or 5 players.

stops eventually because the arms become less and less worthwhile to explore when x goes down, due to the monotone hazard rate in Assumption 1.

Fourth, the cutoff effort level x_I^* for inaction is decreasing with the number of players. When $I = 1$, we are in the special case of a single player, and the cutoff effort x_1^* is higher than the one for multiple players. This makes sense because the shared lumpy cost among I players is lower. Meanwhile, from the perspective of a member in the team, having more members is beneficial in two aspects. One is direct in that the lumpy cost is shared among more players. The other is indirect in that the amount of learning is larger (lower x_I^*) in bigger groups because of information from learning is shared to a greater extent.

Finally, the myopic cutoff effort is x_0. Myopic players consider only the flow payoff. In this model, the flow is $p - x + \frac{Vf(x)}{F(x)}\left(\frac{p-x}{r} - U_I(x) - \frac{L}{I}\right)$, strictly decreasing in V because $U_I(x) \geqslant \frac{p-x}{r}$. Hence myopic players should never search downward because it is always detrimental to flow payoff.

It is interesting to look at the limit of the solution as $\overline{V} \to \infty$. While $\overline{V}$ approaches infinity, it takes less and less time for the team to reach x_I^* and stop. In the limit, it takes no time to achieve the preferred amount of learning. Also, for any fixed x, the value function monotonically converges to

$$\begin{aligned} U_I^\infty(x) &\equiv \lim_{\overline{V}\to\infty} U_I(x) \\ &= \begin{cases} \frac{p-x}{r} + \frac{1}{rF(x)}\int_{x_I^*}^{x}\left(F(s) - \frac{Lr}{I}f(s)\right)ds & \text{if } x \in (x_I^*, x_0] \\ \frac{p-x}{r} & \text{if } x \in (0, x_I^*] \end{cases}, \end{aligned}$$

which is the supremum of value functions for all $\overline{V}$. This limit function is interpreted as the per-capita value achievable if learning takes no time.

Figure 1 summarizes the solution for cooperative problem. The left panel shows the per-capita value functions $U_I^{\infty}(x)$ for $I = 1$ (thick solid curve), $I = 2$ (dashed curve) and $I = 5$ (dot-dash curve). The thin solid line is the payoff of not experimenting at all. The three curves touch the thin solid line at x_1^*, x_2^* and x_5^*, respectively. The right panel shows the time path of x starting from x_0 when $\overline{V} \to \infty$, with 1, 2 or 5 players. After a fast initial decline, they stop at the corresponding cutoff effort levels.

4 Strategic Problem

In a game among $I \geqslant 2$ players, incentives depend on positive information externalities, resulting in very different equilibrium dynamics. In approaching the problem, the HJB of a player is provided and the best response correspondence derived. Properties shared by all stationary Markov equilibria are then available. Since the focus is on symmetric equilibrium, the main result (Theorem 4) solves for the unique symmetric pure strategy stationary Markov equilibrium in closed form, in which learning is too slow compared to social optimum. Some remarks and testable implications follow.

4.1 Best Response Function

For any strategy profile $\{v_i\}_{i=1}^I = \{\nu_i(x)\}_{i=1}^I$, define $W_i(x)$ as the value function for Player i, treating as given the strategy profile of all other players. Let $\nu_{-i}(x) \equiv \sum_{j \neq i} \nu_j(x)$ be the aggregate action contributed by all other players and $\mathcal{V}(x) \equiv \sum_{i=1}^I \nu_i(x)$ be the aggregate action of all players. The HJB for Player i is

$$\begin{aligned} rW_i(x) &= (p-x) + \max_{v_i \in [0,\overline{v}]} v_i \left\{ \frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) - L \right) - W_i'(x) \right\} \\ &\quad + \nu_{-i}(x) \left\{ \frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) \right) - W_i'(x) \right\}. \end{aligned} \tag{7}$$

The first term on the right-hand side is still the flow payoff. The second term (the one with the maximum operator) is decomposed as the loss from breakdown and the benefit from learning generated by Player i's *own* action. The term in the second line arises only in the non-cooperative problem; it consists of loss and benefit brought by *other* players. This

HJB equation differs from the one of the cooperative problem (or single player problem) in an important way, in that the information provided by other players' actions benefits Player i with more learning and lower probability of triggering the breakdown. This type of information externality is exactly "learning from others' (lack of) mistake."

A strategic player who cares about the second term in (7) only faces a problem that is linear in v_i, so the shape of W_i completely determines the best response. The aggregate action from other players, ν_{-i}, affects W_i and thus serves indirectly as the argument of the best response:

$$BR_i(\nu_{-i}) \begin{cases} = \overline{v} & \text{if } \frac{f(x)}{F(x)}\left(\frac{p-x}{r} - W_i(x) - L\right) > W_i'(x), \\ \in [0, \overline{v}] & \text{if } \frac{f(x)}{F(x)}\left(\frac{p-x}{r} - W_i(x) - L\right) = W_i'(x), \\ = 0 & \text{if } \frac{f(x)}{F(x)}\left(\frac{p-x}{r} - W_i(x) - L\right) < W_i'(x). \end{cases} \tag{8}$$

A pure strategy stationary Markov equilibrium requires that $\nu_i(x) \in BR_i(\nu_{-i})$ for all $x \in (0, x_0]$. Proposition 2 states that in any pure strategy stationary Markov equilibrium of an I-player game, the amount of learning must lie between those of a single player problem and an I-player cooperative problem.

Proposition 2 *In any pure strategy stationary Markov equilibrium, (a) $\mathcal{V}(x) > 0$ for $x > x_1^*$, and (b) $\mathcal{V}(x) = 0$ for $x < x_I^*$.*

Proof. See Appendix. ■

4.2 Symmetric Equilibrium

A pure strategy stationary Markov equilibrium is symmetric if $\nu_i(\cdot) = \nu(\cdot)$ for all $i = 1, \ldots, I$. As a result, symmetry implies the value functions satisfy $W_i(\cdot) = W(\cdot)$. The following Proposition solves for the cutoff effort level where the actions switch from positive to zero.

Proposition 3 *In any symmetric pure strategy MPE, $\nu(x) > 0$ if $x > x_1^*$, and $\nu(x) = 0$ if $x < x_1^*$.*

Proof. See Appendix. ■

Proposition 3 indicates that an I-player game leads to the same inaction region $[0, x_1^*]$ as does a single player problem, which is clearly larger than the planner's optimal inaction region $[0, x_I^*]$. In other words, when a single player optimally decides to maintain the current effort level, then adding another player will not prompt her to lower the effort level if we restrict attention to symmetric equilibrium. Here is the intuition. Suppose with multiple players, the cutoff effort level is lowered to $\hat{x} < x_1^*$. Is it sequentially rational for Player i to stick to this stopping rule when facing a state slightly above $\hat{x}$? No. At such a state the learning will stop soon, and the cost of experimentation is approximately $\frac{f(\hat{x})}{F(\hat{x})} v_i L > \frac{f(x_1^*)}{F(x_1^*)} v_i L = \frac{v_i}{r}$, while the benefit from experimentation is approximately $-v_i W_i'(\hat{x}) = \frac{v_i}{r}$. Hence, cost overwhelms benefit, contradicting optimality for Player i.

The cutoff action only informs us that the *eventual amount* of decrease in effort is insufficient in equilibrium, compared to first best. It is also of interest to look at the *speed* of decrease in effort because slow decrease means staying at high effort levels for longer period of time. The following theorem, main result of the section, characterizes the unique symmetric equilibrium of the game, where the decline in effort is indeed slow.

Theorem 4 *Suppose* $\overline{v} \geqslant \frac{\int_{x_1^*}^{x_0} (F(s) - Lrf(s)) ds}{(I-1)L\underline{f}}$. *There exists a unique symmetric pure strategy stationary Markov equilibrium in the I-player strategic problem. Furthermore, the equilibrium features*

$$\nu(x) = \begin{cases} \frac{\int_{x_1^*}^{x} (F(s) - Lrf(s)) ds}{(I-1)Lf(x)} & \textit{if } x \in (x_1^*, x_0] \\ 0 & \textit{if } x \in (0, x_1^*] \end{cases}, \tag{9}$$

$$W(x) = U_1^{\infty}(x). \tag{10}$$

Proof. See Appendix. ■

The condition $\overline{v} \geqslant \frac{\int_{x_1^*}^{x_0} (F(s) - Lrf(s)) ds}{(I-1)L\underline{f}}$ in the theorem is not necessary for existence and uniqueness; it is imposed to avoid corner solutions for $\nu(\cdot)$. There are two noteworthy remarks. First, in the unique symmetric equilibrium, the individual action $v = \nu(x)$ is interior solution, in contrast with the planner's problem. This means the players are indifferent among any action in the range $[0, \overline{v}]$ for $x \in (x_1^*, x_0]$, but choosing the specific action as part of the symmetric equilibrium.

Second, the indifference among actions implies rent dissipation. Given the equilibrium strategies of other players, one can do equally well by choosing the highest action $\overline{v}$. As $\overline{v} \to \infty$, the difference between the payoff of this player in the game and payoff when playing alone vanishes (equation (10)), demonstrating an extreme form of rent dissipation. Although having other players in the game seemingly benefits a player, the procrastination in learning and the resulting high flow cost almost entirely negates the rent.

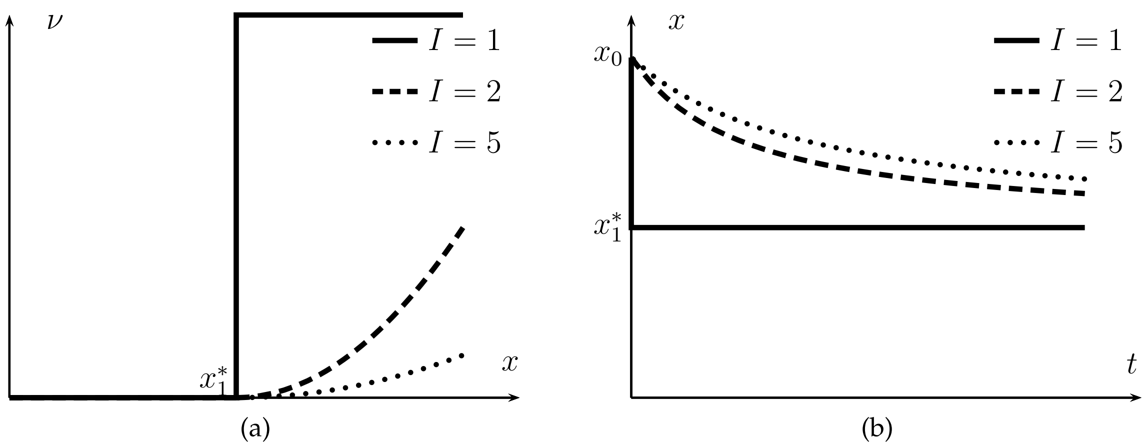

Figure 2: Symmetric Equilibrium. (a) Equilibrium strategy for 1, 2 or 5 players. (b) Time paths for 1, 2 or 5 players.

Having discussed the equilibrium strategies, it is interesting to elaborate on the behavior of the time path of effort. Absent any breakdown, the path $x(\cdot)$ is uniquely determined by the differential equation (2). From Theorem 4, we know that x decreases over time as long as $x > x_1^*$, and will never reach a state below x_1^*. Moreover, as is evident from (9), the individual action (and also the aggregate action) shrinks to zero when $x \downarrow x_1^*$, reflecting a severe downscaling in actions when the effort level is close to the cutoff. It is shown below in Proposition 5 that the state never exactly reaches x_1^*.

Proposition 5 *When $c \leqslant x_1^*$, the time path $x(t)$ satisfies (a) $\lim_{t\to\infty} x(t) = x_1^*$; (b) $x(t) > x_1^*$ for all $t \in \mathbb{R}_+$.*

Proof. First, note that by Assumption 3, f defined on $[0, x_1^*]$ must have an upper bound $\overline{f} > 0$.

Part (a): $x(t)$ is non-increasing, and is bounded below by x_1^*. If $\lim_{t\to\infty} x(t) = \hat{x} > x_1^*$, then by Assumption 1, $x'(t) = -I\nu(x) < -\frac{I\int_{x_1^*}^{\hat{x}}[F(s)-Lrf(s)]ds}{(I-1)L\overline{f}} < 0$ for all $t \in \mathbb{R}_+$, a contradiction.

Part (b): $\nu(x)$ is bounded above by $\hat{\nu}(x) \equiv \frac{\int_{x_1^*}^{x}(\overline{f}+Lr\kappa)(s-x_1^*)ds}{(I-1)L\underline{f}}$. The path of x, when ν is replaced by $\hat{\nu}$, is always above x_1^*, so it must be true for the original path too. ■

The proposition says that without breakdowns, the long-run limit of x is x_1^*, so that the probability of eventually learning the threshold is $1 - F(x_1^*)$. It also implies that for threshold realizations slightly above x_1^*, the time to learn its exact location becomes unbounded. Actually, we can derive the speed of convergence in the following.

Proposition 6 *Suppose $c \leqslant x_1^*$. If $\lim_{x \downarrow x_1^*} \frac{d}{dx}\left(\frac{F(x)}{f(x)}\right) = b > 0$, then the time path $x(t)$ converges to x_1^* at speed t^{-1}.*

Proof. See Appendix. ■

Figure 2 shows the equilibrium strategy $v_i = \nu_i(x)$ and the time path, for games with $I = 2, 5$ players. As a comparison, the time path for the single player problem is also presented.

4.3 Comparative Statics and Implications

The dynamics of the game depend on the primitives. Fix $x_0 > 0$, we first examine the effect on the inaction region of changing the discount rate, the lumpy cost and the prior distribution function.

Corollary 7 *The cutoff effort level x_1^* is increasing in L and r. Moreover, if two prior distribution functions F and F' are hazard rate ordered such that $\frac{f(x)}{F(x)} < \frac{f'(x)}{F'(x)}$ for all $x \in [0, x_0]$, then x_1^* under F is lower than that under F'.*

Proof. By Assumption 1, the above comparative statics are immediate from the equation $\frac{F(x_1^*)}{f(x_1^*)} = Lr$. ■

The cutoff is directly related to the probability of eventual learning $1 - F(x_1^*)$. Hence, projects with less costly breakdown operated by a more patient group will likely to have a lower cutoff effort level and higher probability of eventual learning. As players become perfectly patient, sufficient learning is achieved. The second part of Corollary 7 says that if the prior distribution is more favorable (concentrated more in the lower end), then the cutoff is lower.

The model also generates some comparative statics regarding the equilibrium actions. However, these results have testable implications only if the action history is also available to an outside econometrician.

There are contexts where the action/effort history is observable. Examples include leverage levels in financial sectors, amount of pollution to the environment, etc. In these situations, we have the following predictions.

Lemma 8 *Fix distribution F. At any state $x > x_1^*$, the individual action $\nu(x)$ (as well as the aggregate action $\mathcal{V}(x)$) is decreasing in L, r, and I.*

Proof. See Appendix. ■

The lemma implies that (a) in scenarios with more severe consequence from breakdown, players are less eager to bring down the effort level; (b) the same is true if the group of players is less patient; (c) the aggregate action (the speed of decline in effort) is decreasing in the number of players I. While statement (a) is somewhat expected, (b) and (c) deserve short comments. Statement (b) is surprising at first glance if one fits the story into a "quality maintenance" context. Why should less patient players keep a higher quality (effort) level? There are two reasons. First, unlike in repeated games, the lumpy cost of breakdown occurs immediately when $x < c$ so that there is no discounting in the size of the "punishment". Second, there is learning in the threshold and it is precisely the impatient players who do not wish to learn more at the cost of current payoff. Statement (c) implies that the negative impact of free-riding outweighs the number of players, so that even the aggregate action is decreasing in I. As $I \to \infty$ the aggregate action becomes half of that in a two-player game.

There are even more situations where the action/effort history is unobservable or difficult to quantify. For instance, product qualities from manufacturers are unobservable, and maintenance effort levels for power plants are hard to measure. Thus one cannot use the action or effort path to make predictions. However, there is an implication based on the hazard rate: the hazard rate of breakdown is decreasing over time. It is useful for an outsider econometrician as long as the occurrence of incidence (breakdown) is observable.

Lemma 9 *The hazard rate of breakdown, $I\nu(x)\frac{f(x)}{F(x)}$, is decreasing over time.*

Proof. To see this, note that the hazard rate of breakdown is

$$I\nu(x)\frac{f(x)}{F(x)} = \frac{I\int_{x_1^*}^{x}(F(s) - Lrf(s))ds}{(I-1)LF(x)},$$

and this is increasing in x because

$$\begin{aligned} & \frac{d}{dx}\left(I\nu(x)\frac{f(x)}{F(x)}\right) > 0 \\ \Leftrightarrow \quad & \left(\frac{F(x)}{f(x)} - Lr\right)F(x) > \int_{x_1^*}^{x}\left(\frac{F(s)}{f(s)} - Lr\right)f(s)ds, \end{aligned}$$

which is true by Assumption 1. ■

Hence, as time goes by, x decreases, and the hazard rate of triggering a breakdown is decreasing. This coincides with the general observation that long-established firms tend to maintain a steady quality and less likely suffer scandals of quality issues.

5 Discrete Time and Convergence

This section studies a discrete-time version of the game. In particular, the game features alternating moves where $I \geqslant 2$ players take turns to change the current effort to a possibly new level. A breakdown arrives as soon as the effort level falls below the unknown threshold after some player's move, and that player alone bears the lumpy cost from the breakdown. The breakdown also ends the game and gives a terminal payoff, just like the continuous time version.

The purpose of this section is to let the frequency of moves go to infinity (i.e., period length goes to zero) and to show that the realized time path of effort in the discrete-time game converges to that of the continuous-time version. In this sense, the outcome of the continuous-time game is robust to perturbations in the fineness of time grids.

Formally, time is discrete and the time periods are indexed by $n = 1, 2, \ldots$. The horizon of the game is random. The period length is $\Delta > 0$. For simplicity, let there be $I = 2$ players labeled $i = 1, 2$. In each of the odd periods n, Player 1 chooses an action $v_1(n) \in [0, \overline{v}]$ (again, $\overline{v} > 0$), provided the game has not ended. In even periods, Player 2 chooses an action $v_2(n) \in [0, \overline{v}]$. The effort level x evolves as follows:

$$x(n) = \begin{cases} x(n-1) - v_1(n)\Delta & \text{for odd } n, \\ x(n-1) - v_2(n)\Delta & \text{for even } n, \end{cases}$$

where $x(0)$ is defined to be some $x_0 > 0$, given at the beginning of Period 1. Hence, in each period the effort decreases by $v_i(n)\Delta$, where the identity of i depends on who makes the move in that period. Note that $v_i(n)$ is similarly interpreted as the "speed of decline," but its magnitude is twice as large as the counterpart in continuous time, since here each player only contributes in half of the time.

As in the continuous time model, the unknown threshold $c \in [0, x_0)$ has c.d.f. $F(\cdot)$ that satisfies Assumptions 1 and 2. Moreover, the discreteness demands a slightly stronger requirement on the hazard rate:

Assumption 4 *(Strongly Monotone Hazard Rate)*
The inverse hazard rate $\frac{F(\cdot)}{f(\cdot)}$ *is strongly increasing, i.e., for some* $\underline{b} > 0$, $\frac{F(x)}{f(x)} - \frac{F(y)}{f(y)} \geqslant \underline{b}(x - y)$ *for* $0 \leqslant y \leqslant x \leqslant x_0$.

Given the evolution of x, the game terminates at the end of Period $\overline{n}$ if $\overline{n}$ is the smallest $n \geqslant 1$ such that $x(n) \leqslant c$, and $\overline{n}$ can be infinity. Differently from the continuous time version, even the threshold is triggered in Period $\overline{n}$, it is not perfectly inferred; players only know that $c \in [x(\overline{n}), x(\overline{n}-1))$.

Payoffs come in flows and lumps. Let $r > 0$ be the common real time discount rate of the players, and $\delta \equiv e^{-r\Delta}$ be the discount factor between periods. Define for convenience that $\tilde{\Delta} \equiv \frac{1-e^{-r\Delta}}{r} = \frac{1-\delta}{r}$. In each Period $n \leqslant \overline{n}$, the flow benefit is $\int_0^\Delta e^{-rs} p ds = p\tilde{\Delta}$ for both players. The flow cost from effort is $\int_0^\Delta e^{-rs} x(n) ds = x(n)\tilde{\Delta}$ for the player who moves in Period n, and is $x(n-1)\tilde{\Delta}$ for the other.[8] If the game ever ends with terminal period $\overline{n}$, then the player who moves in that period bears the lumpy cost $L > 0$. Moreover, each player receives a terminal lump sum $\frac{p-x(\overline{n}-1)}{r}$ equalling the present discounted value of a flow $(p - x(\overline{n}-1))\tilde{\Delta}$ per period from then on.[9]

Recall that in the continuous-time model, the lumpy cost for Player i at the end of the game is $L\frac{v_i(\overline{t})}{V(\overline{t})}$. This can be seen as a limit payoff in the discrete-time game when period length goes to zero. Consider two consecutive periods n and $n+1$, where Player i moves in the former and Player $-i$ moves in the latter. Conditional on the fact that the game ends within the time window $[(n-1)\Delta, (n+1)\Delta]$, it ends in Player i's turn with probability $\frac{v_i(n)\Delta}{(v_i(n)+v_{-i}(n+1))\Delta} = \frac{v_i(n)}{v_i(n)+v_{-i}(n+1)}$. Once we restrict attention to Markov strategies (see below) that are piecewise Lipschitz continuous in x, this ratio converges to the share of loss in the continuous time model as $\Delta \to 0$.

A pure Markov perfect strategy $\nu_i(\cdot)$ is a mapping from the state space $[0, x_0]$ into the action space $[0, \overline{v}]$, for $i = 1, 2$. Again, we are interested in symmetric MPE in pure strategies. Denote it as $\nu_1(\cdot) = \nu_2(\cdot) = \nu(\cdot)$.

5.1 Existence and Convergence

Having defined the discrete-time model, I present the existence result below, followed by the convergence result that links the discrete- and continuous-time models.

Proposition 10 *Fix a $\Delta > 0$ that is small enough. Suppose $\overline{v} \geqslant 2\frac{\int_{x_1^*}^{x_0}(F(s)-Lrf(s))ds}{L\underline{f}}$. There exists*

[8]This difference in cost results from the asynchronous moves. The difference is unimportant, and it vanishes in continuous time model. However, we cannot call $x(n)$ the common effort level. Instead, we call it the frontier of effort.

[9]Here the imaginary "continuation" payoff is based on the assumption that both players revert to the effort in the last period as the cheapest safe action. Actually, if we allow the game to continue, then this reversion is indeed optimal for small enough Δ.

a continuum of symmetric pure strategy MPE $E(y_0; \Delta)$, parametrized by $y_0 \in [\underline{y}_0, x_0]$ for some $\underline{y}_0 \in (x_1^*, x_0)$, of the following form:

$$\nu(x) = \begin{cases} \frac{x - y_{k+1}}{\Delta} & \text{if } x \in (y_{k+1}, y_k] \text{ for some } k \geqslant -1 \\ 0 & \text{if } x \leqslant x_1^* \end{cases}$$

where $y_{-1} \equiv x_0$, $x_1^* < \ldots < y_{k+1} < y_k < \ldots < y_1 < y_0$ and $\lim_{k\to\infty} y_k = x_1^*$.

Proof. See Appendix. ■

Proposition 10 has two facets. First, it claims that there exists some symmetric pure strategy MPE with the *skimming property*, in which there is a strictly decreasing sequence of "critical levels" $\{y_k\}_{k=1}^{\infty}$ of the state variable such that the player currently making a move always chooses an action bringing the state x down to the highest critical level strictly below x. The sequence starts with $y_0 \in [\underline{y}_0, x_0]$ and monotonically converges to x_1^*, so on path Player 1 brings the state down from x_0 to y_0, then Player 2 brings it to y_1, and again Player 1 sets the new state at y_2, etc. This process goes all the way down to x_1^* asymptotically. In equilibrium the current mover facing state y_k is indifferent between staying at y_k and moving down to y_{k+1}, although the equilibrium requires her to choose the latter.

Second, it also states that there is some indeterminacy leading to multiple equilibria. As is usually the case for equilibrium with the skimming property (bargaining problem for instance), multiplicity arises because unlike other critical levels of the state variable, the initial state may not satisfy the indifference condition. As a result, there is leeway in choosing $y_0 \in [\underline{y}_0, x_0]$ as the first critical point satisfying the indifference condition. Once y_0 is chosen, everything else is determined.

The indeterminacy does not pose a problem in the limit as $\Delta \to 0$, as we will show in Proposition 11. The main idea is that with smaller Δ, the sequence $\{y_k\}_{k=0}^{\infty}$ is denser, and the interval $[\underline{y}_1, x_0]$ is narrower. Hence, even there is a continuum of equilibria, they differ by less and less as $\Delta \to 0$.

The next proposition shows the convergence of time path as $\Delta \to 0$. Thanks to Proposition 10, for any small enough period length $\Delta > 0$, we can select a symmetric pure strategy MPE $E(y_0(\Delta); \Delta)$ indexed by $y_0(\Delta) \in [\underline{y}_0(\Delta), x_0]$, and the resulting path is denoted by $x(\cdot; \Delta)$. Hence, if a sequence $\Delta_n \downarrow 0$, then there exists $N > 0$ s.t. for all $n \geqslant N$, the selection $E(y_0(\Delta_n); \Delta_n)$ exists, generating a sequence of time paths $\{x(\cdot; \Delta_n)\}_{n=N}^{\infty}$.

Proposition 11 *Suppose* $\overline{v} \geqslant 2\frac{\int_{x_1^*}^{x_0}(F(s)-Lrf(s))ds}{L\underline{f}}$. *For any fixed sequence* $\Delta_n \downarrow 0$ *there exists* $N > 0$ *s.t. the sequence of paths* $\{x(\cdot;\Delta_n)\}_{n=N}^{\infty}$ *generated by any selection converges to that from the continuous time model at every* $t \in [0, \overline{t})$.

Proof. See Appendix. ■

Hence, viewed in real time, the path of effort levels in the discrete time model is a step function, but it converges to a decreasing smooth function, which is the outcome of the continuous-time model. To some extent, this convergence serves as a robustness check to the main model, assuring that both the setup and the equilibrium predictions are valid as a proper limit of some discrete time model.

Convergence does not always hold for any assumptions on the payoff structure. If, for example, the negative impact of breakdown does not involve a lump sum L but instead costs the entire terminal benefit as an "endogenous punishment", then there are no equilibrium outcomes of the discrete time version in the vicinity of the continuous time outcome.

6 Extensions

This section returns to the continuous-time model, but considers several branches of extensions. First, we aim to investigate the dependence of equilibrium on the monitoring structure. Second, we relax the assumption of a constant lump sum L across states, allowing for dependence of L on x. Finally, we use the ironing approach to solve the problem when Assumption 1 fails.

6.1 Unobservable Actions

In this extension I consider an alternative monitoring structure: unobservable actions with public breakdown. The benchmark case where both the action history and the breakdown are observable is analyzed in Section 4, while the other extreme—unobservable actions and breakdown—is equivalent to a single player problem because no information is transmitted whatsoever. The monitoring structure considered here is a compromise; players secretly operate their own learning enterprises with a common threshold, leading to potentially different effort paths over time, but the event of someone triggering the breakdown is instantly revealed. The relevance of this variation is justified by the observation that it is easier to get notified of a breakdown than to watch the operation of one's opponent. Information externalities still persist, although in a less straightforward way.

Formally, there are two players $i = 1, 2$. Time is continuous. For $i = 1, 2$, there is a (potentially different) stopping time $\bar{t}_i > 0$, which is specified later. The horizon of the game, $\bar{t}$, is then defined as $\bar{t} \equiv \max\{\bar{t}_1, \bar{t}_2\}$. At each time $t \in [0, \bar{t}_i]$, Player i takes private action $v_i(t) \in [0, \bar{v}]$. Each player has an *individual* effort level $x_i > 0$ the law of motion of which is

$$x_i(t) = x_0 - \int_0^t v_i(s) ds, \tag{11}$$

for some common initial value $x_0 > 0$. The effort level is no longer common because of the lack of observability of the other player's action.

The unknown threshold $c \in [0, x_0)$ is common, the distribution of which satisfies the same conditions as in the main model. For $i = 1, 2$, define the stopping time as $\bar{t}_i \equiv \inf\{t : x_i(t) \leqslant c\} \in \mathbb{R}_+ \cup \{+\infty\}$. Hence, Player i's part in the game ends as soon as $x_i(t) \leqslant c$. If some player triggers the threshold first, then this event becomes public immediately without revealing the effort level that pulls the trigger.

Payoffs come in flows and lumps. Before time $\bar{t}_i$, the flow payoff for Player i is $p - x_i$: fixed benefit minus the cost of effort. If $\bar{t}_i < \infty$, then her part of the game ends at $\bar{t}_i$, with a terminal cost L from breakdown and a terminal benefit $\frac{p-c}{r}$. If Player i's part has ended but Player j's has not, then Player j moves on as a single player.

At time $t \in [0, \bar{t}_i]$, the information available to Player i is her private action history $\{v_i(s)\}_{s<t}$ as well as the time at which Player j hits the threshold, if at all. Since time itself contains information, the payoff-relevant states are individual effort level, time, and possibly the time at which the opponent ends her part. A pure Markov strategy of Player i is thus a pair of mappings

$$\begin{aligned} &\nu_i^0(x_i, t) : [0, x_0] \times \mathbb{R}_+ \to [0, \bar{v}], \\ &\nu_i^1(x_i, t, \bar{t}_j) : [0, x_0] \times \mathbb{R}_+ \times \mathbb{R}_+ \to [0, \bar{v}], \end{aligned}$$

where ν_i^0 and ν_i^1 describe the pure strategy of Player i before and after her opponent hits the threshold, respectively. Let $\mathcal{N}_i$ be the set of pure Markov strategy pairs where both ν_i^0 and ν_i^1 are piecewise Lipschitz continuous in x_i. The first component of a pure strategy,

ν_i^0, uniquely defines a path of effort $x_i(t)$ from the differential equation

$$\frac{dx_i}{dt} = -\nu_i^0(x_i, t), \;\; x_i(0) = x_0, \tag{12}$$

when $\min\{\bar{t}_i, \bar{t}_j\}$ has not arrived. In what follows, *path of effort* refers to the effort level as a function of time conditional on the event of no breakdown so far. The second component, ν_i^1, is simply the single-player policy in the continuation game where the posterior distribution of the threshold is obtained by Bayes-updating.

A pure strategy Markov equilibrium is a profile of pairs $\{(\nu_1^0, \nu_1^1), (\nu_2^0, \nu_2^1)\}$ such that (ν_i^0, ν_i^1) constitutes a best response to (ν_j^0, ν_j^1) for any feasible combination of (x_i, t) or $(x_i, t, \bar{t}_j)$. As we will soon see, the game does not admit a pure strategy Markov equilibrium, so it is necessary to consider mixed Markov strategies $\sigma_i : [0,1] \to \mathcal{N}_i$ for $i = 1, 2$. A mixed strategy Markov equilibrium is a profile (σ_1, σ_2) such that any (ν_i^0, ν_i^1) in the support of σ_i is a best response to σ_j for any feasible combination of (x_i, t) or $(x_i, t, \bar{t}_j)$. Given a mixed strategy σ_i of Player i and conditional on no breakdown, denote $P_i(t|x) \equiv \Pr(s \leqslant t | x_i(s) = x)$ as the cumulative distribution of time conditional on Player i's effort level being x. Also, define $x_i^h(t) \equiv \inf\{x : P_i(t|x) = 1\}$ and $x_i^l(t) \equiv \sup\{x : P_i(t|x) = 0\}$. Obviously $x_i^h(t) \geqslant x_i^l(t)$ for all $t \in \mathbb{R}_+$. Moreover, x_i^h and x_i^l are both non-increasing. The following lemma makes comparison of payoffs in extreme cases.

Lemma 12 *Fix Player j's (possibly mixed) strategy σ_j. For Player i, consider two paths of effort $x_i(\cdot)$ and $\tilde{x}_i(\cdot)$ resulting from two pure strategies, conditional on no breakdown.*

(i) If for some $0 \leqslant t_a < t_b \leqslant \infty$ we have $x_i(t) > \tilde{x}_i(t) > x_j^h(t)$ for $t \in (t_a, t_b)$ and $x_i(t) = \tilde{x}_i(t)$ otherwise, then Player i strictly prefers $\tilde{x}_i$ to x_i.

(ii) If for some $0 \leqslant t_a < t_b \leqslant \infty$ we have $x_j^l(t) > x_i(t) > \tilde{x}_i(t) \geqslant x_1^*$ for $t \in (t_a, t_b)$ and $x_i(t) = \tilde{x}_i(t)$ otherwise, then Player i strictly prefers $\tilde{x}_i$ to x_i.

Proof. See Appendix. ■

The lemma says that on one hand, if two paths are both slow in decline so that Player j has a lower effort level at all times, then Player i prefers the lower path. On the other hand, if two paths (always above x_1^*) are both quick in decline so that Player j has a higher effort level at all times, then Player i also prefers the lower path. Intuitively, when the choice of paths does not alter the probability of incurring the lumpy cost, faster decline in effort is always preferable because of the savings in flow cost.

Proposition 13 utilizes the previous lemma to rule out the existence of any pure strategy equilibrium.

Proposition 13 *There is no pure strategy Markov equilibrium.*

Proof. See Appendix. ■

An equilibrium is said to have *deterring property* if for $i = 1, 2$, σ_i puts zero probability on all strategies with $\nu_i^1 > 0$, i.e. once a player hits the threshold first, the other one immediately stops lowering effort. I focus on symmetric Markov equilibria with this deterring property. Denote $W_i(x_i, t)$ as Player i's value function before anyone hits the threshold. Denote $P_i(t|x) \equiv \Pr(s \leqslant t|x_i(s) = x)$ as the distribution of time conditional on Player i's effort level being x, generated by her equilibrium mixed strategy σ_i. From Player i's point of view, the equilibrium distribution $P_j(t|x)$ is given, and her HJB is:

$$\begin{aligned} rW_i(x_i, t) &= (p - x_i) + \max_{v_i \in [0, \overline{v}]} v_i \left\{ \frac{Q_{jx}(x_i, t)}{Q_j(x_i, t)} \left(\frac{p - x_i}{r} - L - W_i(x_i, t) \right) - W_{ix}(x_i, t) \right\} \\ &\quad - \frac{Q_{jt}(x_i, t)}{Q_j(x_i, t)} \left(\frac{p - x_i}{r} - W_i(x_i, t) \right) + W_{it}(x_i, t), \end{aligned} \tag{13}$$

where $Q_j(x_i, t) \equiv \int_0^{x_i} [1 - P_j(t|s)] f(s) ds$ is the total probability that $c < x_i$ and Player j reaches c after time t. In other words, $Q_j(x_i, t)$ is the unconditional "real threat" of the lumpy cost that Player i faces, and it is smaller than $F(x_i)$ because Player j might hit the threshold before Player i. The subscripts x and t denote the partial derivatives. The right-hand side consists of several parts. The first term is the net flow payoff as usual. The second term captures two effects of taking action v_i. One is the cost of triggering the breakdown herself, and the other is the benefit from *learning through state*. The third term is indirect learning from the opponent's lack of breakdown, equalling the conditional probability rate that Player j hits the threshold multiplied by the change of payoff in that case. The last term is the capital gain from changes in time. The last two terms are both benefits from *learning through time*. They are novel to this unobservable actions case because the information externality from the other player comes indirectly through time.

Formally, for $x > x_1^*$ define $T(x) \equiv \frac{x_0 - x}{\overline{v}} + \frac{1}{r} \ln \xi(x)$ where $\xi(x) \equiv \left[1 - e^{-\frac{x - x_1^*}{Lr}} \frac{f(x)}{f(x_1^*)} \right]^{-1}$. Note that $\xi(x) > 1$, $T(x) > \frac{x_0 - x}{\overline{v}} \geqslant 0$, and $-\frac{1}{\overline{v}} < T'(x) < 0$ for $x \in (x_1^*, x_0]$. Consider a class of pure Markov strategies $\overline{\mathcal{N}} \equiv \{\nu^0(x,t;t_0) : t_0 \in [0, T(x_0)]\}$ where

$$\nu^0(x,t;t_0) \equiv \begin{cases} \frac{f(x)Lr^2}{[f(x)-Lrf'(x)]\left(1-e^{-r[t-(x_0-x)/\overline{v}]}\right)\xi(x)+f(x)Lr^2/\overline{v}} & \text{if } x \geqslant x_1^*, t_0 \leqslant t \leqslant T(x) \\ \overline{v} & \text{if } x \geqslant x_1^*, t > T(x) \\ 0 & \text{otherwise} \end{cases} . \tag{14}$$

The strategy $\nu^0(x,t;t_0)$, indexed by t_0, stipulates a player to remain inactive for some initial period $[0, t_0]$. If we integrate out the differential equation (12) using (14), then for each $t_0 \in [0, T(x_0)]$ there is a weakly decreasing time path $x(\cdot; t_0)$ implicitly defined by the equation

$$\left(e^{r[t-(x_0-x)/\overline{v}]} - 1\right)(\xi(x) - 1)^{-1} = \left(e^{rt_0} - 1\right)(\xi(x_0) - 1)^{-1}$$

The following proposition describes a symmetric mixed strategy Markov equilibrium with deterring property, where the support of mixed strategies is exactly $\overline{\mathcal{N}}$.

Proposition 14 *The following strategy σ describes a symmetric mixed strategy Markov equilibrium with deterring property.*

$$\begin{aligned} &\Pr_\sigma(\nu^0(\cdot,\cdot;t_0) : t_0 \leqslant \tau) = (e^{r\tau} - 1)(\xi(x_0) - 1)^{-1}, \qquad (15) \\ &\nu^1(x,t,\bar{t}_j) \equiv 0. \end{aligned}$$

Furthermore, absent any breakdown, the time path corresponding to any index $t_0 \in [0, T(x_0)]$ converges to x_1^*.

Proof. See Appendix. ■

Actually, we can solve for the value function in closed form:

$$W(x,t) = \begin{cases} \frac{p-x}{r} + \frac{1}{rQ(x,t)} \int_{x_1^*}^{x} [Q(s,t) - LrQ_x(s,t)]ds & \text{if } x \in (x_1^*, x_0] \\ \frac{p-x}{r} & \text{if } x \in (0, x_1^*] \end{cases} . \tag{16}$$

where

$$Q(x,t) = \int_0^x [1 - P(t|s)]dF(s), \tag{17}$$

$$P(t|x) = \begin{cases} \left(e^{r[t-(x_0-x)/\overline{v}]} - 1\right)(\xi(x) - 1)^{-1} & \text{if } x > x_1^*, t \in \left[\frac{x_0 - x}{\overline{v}}, T(x)\right] \\ 1 & \text{if } x > x_1^*, t \in (T(x), \infty) \\ 0 & \text{otherwise} \end{cases} . \tag{18}$$

The proposition has many implications. First, the equilibrium entails symmetric randomization over a compact set of pure strategies (see (15)). Each such strategy is a member of $\overline{\mathcal{N}}$, indexed by the initial hibernation period $t_0 \in [0, T(x_0)]$. Hence, with the realization of t_0, a player waits at the starting effort level x_0 until t_0 and then takes positive actions resulting in a deterministically decreasing path in effort. All choices of t_0 give the player the same ex ante payoff, and $t_0 > T(x_0)$ leads to lower payoff, so the player is willing to randomize with the particular distribution.

Second, the players are willing to follow the prescribed deterministic path of effort once t_0 is chosen. Taking (16) to the HJB (13), we know that in equilibrium the multiplier on v_i is zero, so Player i is indifferent in choosing any $v_i \in [0, \overline{v}]$. In particular, she does well with the prescribed path. This indifference among all effort levels at any time makes it sufficient to consider only initial randomization on t_0.

Third, the randomization in t_0 means a compact spread of t at which x_i reaches any level between x_0 and x_1^* (see (18)), and conversely a compact spread of x_i at any time t. For the lowest realization $t_0 = 0$, the law of motion (11) yields the effort path of a single player problem: fastest decrease before reaching x_1^* and then stay there. For higher t_0, the decrease in x_i is slower.

Fourth, the game with unobservable actions leads to the same cutoff effort level as in the main model (see (14)). This is not too surprising, considering the fact that both no-information (single player) game and full-information game have the same cutoff effort.

Fifth, the ex ante payoff of a player is the same as in the observable action case, i.e. $W(x_0, 0) = W(x_0)$. This means the observability of action affects only the dynamics of the play, but not the payoffs. Indeed, in the main model the symmetric equilibrium is *interior* so that the value function must assume a particular form to support the indifference; a player is just as willing to choose $\nu = \overline{v}$ any time, which is the single player policy. In the extension here the symmetric equilibrium is in *mixed strategies* so that similar requirements are put on the value function to maintain the willingness to randomize; the

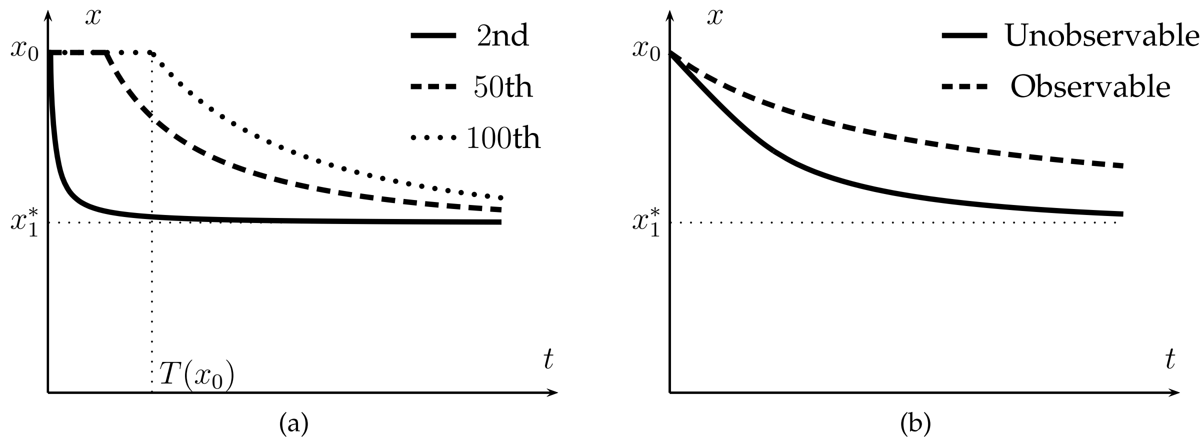

Figure 3: Unobservable actions. (a) Time paths from mixed strategies. (b) Comparison of speed of decline in effort.

single-player path of effort is in the support of the mixed strategy. That is why the payoffs are equal. Again, we have rent dissipation: having other players in the game does not benefit the existing player.

Proposition 15 *Suppose $c \leqslant x_1^*$. If $\lim_{x \downarrow x_1^*} \frac{d}{dx}\left(\frac{F(x)}{f(x)}\right) = b > 0$, then the time path $x(t)$ converges to x_1^* at speed e^{-rt}.*

Proof. See Appendix. ■

Define $T^{\infty}(x) = \lim_{\overline{v} \to \infty} T(x)$. By definition, $(T^{\infty})^{-1}(\cdot)$ is the path of effort following the highest realization of t_0. As $\overline{v} \to \infty$, the pair (t, x) on path fills the entire set $\{(t, x) : x \in [x_1^*, x_0], t \in [0, T^{\infty}(x)]\}$. Panel (a) of Figure 3 plots in the $t - x$ plane the deterministic paths of effort following the 2nd, 50th, and 100th percentile of t_0.

Panel (b) of Figure 3 gives a comparison between the observable and unobservable action cases, for the time paths of the expectation of the lower action conditional on no breakdown. Intuitively, the expected speed of decline in x is faster in the unobservable action case, but it is in turn slower than the no information case (single player case).

6.2 Variable Lumpy Cost

It has been previously assumed that the lumpy cost L is a fixed amount. However, one can readily argue that the lumpy cost can vary with the realized threshold. For example, in the maintenance of environment or health, the realized threshold is usually negatively correlated with the cost upon triggering. On the other hand, in international relations, a high threshold is often associated with a high lumpy cost, when both are indicators of the "toughness" of the country.

In order to obtain well-behaved solutions, we need modify Assumptions 1 through 3:

Assumption 5 *(Modified Monotone Hazard Rate)*
The ratio $\frac{F(\cdot)}{f(\cdot)L(\cdot)}$ *is strictly increasing.*

Assumption 6 *(Strongly Positive Density and Cost)*
Both $f(\cdot)$ *and* $L(\cdot)$ *are uniformly bounded away from zero, i.e.* $\exists \underline{f} > 0$ *s.t.* $\min\{f(c), L(c)\} \geqslant \underline{f}$ *for all* $c \in [0, x_0]$*.*

Assumption 7 *(Lipschitz Continuous Density and Cost)*
Both $f(\cdot)$ *and* $L(\cdot)$ *are Lipschitz continuous, i.e.* $\exists \kappa > 0$ *s.t.* $\max\{|f(x) - f(y)|, |L(x) - L(y)|\} \leqslant \kappa|x - y|$ *for all* $x, y \in [0, x_0]$*.*

Accordingly, we need to define $\tilde{x}_1^*$ as the unique solution to $\frac{F(x)}{f(x)L(x)} = r$. The following result is a counterpart of Theorem 4.

Proposition 16 *Suppose* $\overline{v} \geqslant \frac{\int_{\tilde{x}_1^*}^{x_0} (F(s) - L(s) r f(s)) ds}{(I-1)\underline{f}^2}$*. There exists a unique symmetric pure strategy stationary Markov equilibrium in the I-player strategic problem. Furthermore, the equilibrium features*

$$\nu(x) = \begin{cases} \frac{\int_{\tilde{x}_1^*}^{x} [F(s) - L(s) r f(s)] ds}{(I-1) L(x) f(x)} & \text{if } x \in (\tilde{x}_1^*, x_0] \\ 0 & \text{if } x \in (0, \tilde{x}_1^*] \end{cases}$$
$$W(x) = \begin{cases} \frac{p-x}{r} + \frac{1}{rF(x)} \int_{\tilde{x}_1^*}^{x} (F(s) - L(s) r f(s))\, ds & \text{if } x \in (\tilde{x}_1^*, x_0] \\ \frac{p-x}{r} & \text{if } x \in (0, \tilde{x}_1^*] \end{cases},$$

The proof of this proposition is similar to that of Theorem 4.

6.3 Non-Monotone Hazard Rate: Ironing

That $\frac{F(x)}{f(x)}$ is strictly increasing is assumed for technical simplicity. However, in some real situations the distribution function does not necessarily satisfy this condition. For example, the players may know that the subject of experimentation has two major types: tough or delicate, and within each type there are some noises determining the actual

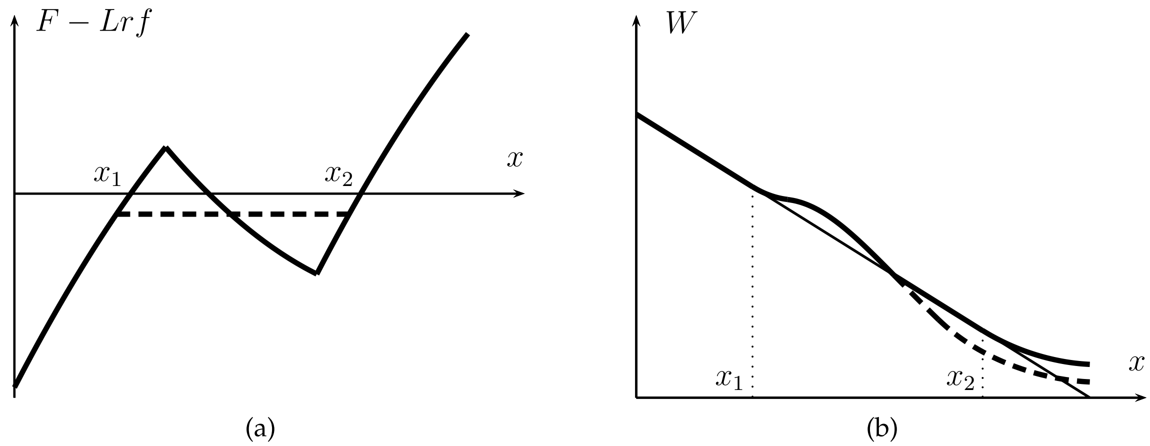

Figure 4: Two candidates. x_1 not worth reaching.

realization of the threshold. In this way the distribution is bimodal, and Assumption 1 may fail. This subsection provides an ironing method to overcome this problem.

In this subsection we discard both Assumption 1 and the requirement that $\frac{F(x_0)}{f(x_0)} \geqslant Lr$. Due to the failure of monotone hazard rate, the solutions to $\frac{F(x)}{f(x)} = Lr$ may not be uniquely determined, or may fail to exist when $\frac{F(x_0)}{f(x_0)} < Lr$. The left panel of Figure 4 depicts a situation where $F(x) - Lrf(x)$ crosses 0 from below at both x_1 and x_2, but the total area under the curve from x_1 to x_2 is negative. After ironing, the curve to the left of x_2 becomes negative everywhere, meaning that x_1 is a wrong candidate for settling point. The right panel shows the hypothetical payoff function if the players stop learning at x_1. The payoff at x_2 goes below the "no move" payoff $\frac{p-x_2}{r}$, implying suboptimality to stop at x_1. Intuitively, the distribution F is such that going downward from x_2, the players have to experience some hardship due to locally higher density f, but ease follows when they do overcome the peak in f. In this case, the forward-looking players choose to stop at x_2 because the cost outweighs the benefit. On the contrary, Figure 5 shows a situation where the ironed curve crosses 0 at x_1, meaning that the local hardship is worth investing for a reward to the far-left.

Formally, define $K(x) \equiv \int_0^x [F(s) - Lrf(s)]ds$, and following the similar trick to Myerson (1981), define $\overline{K}(\cdot) \equiv \text{Vex}(K(\cdot))$ as the convexification of $K(\cdot)$. With this transformation, $\overline{K}'(x)$ is weakly increasing. The result below shows the selection criterion when facing non-monotone hazard rate.

Proposition 17 *(i) If $\overline{K}'(x_0) \leqslant 0$, then any player in the symmetric equilibrium should stay at x_0 forever.*

(ii) If $\overline{K}'(x_0) > 0$ and $\overline{K}'(x)$ has a unique intersection with 0 at $\overline{x}^*$, then it is the cutoff effort

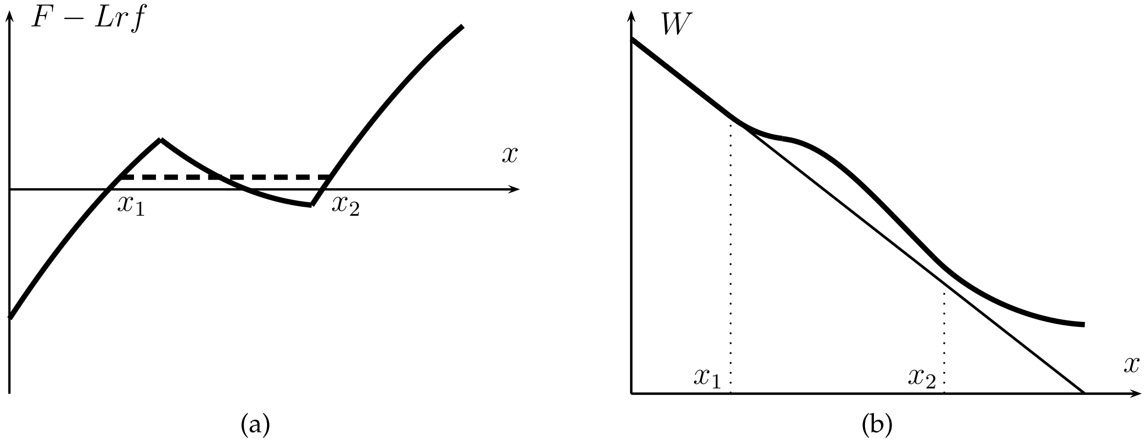

Figure 5: Two candidates. x_1 worth reaching.

where learning stops eventually.

(iii) If $\overline{K}'(x_0) > 0$ and $\overline{K}'(x)$ equals 0 on some interval $[\overline{x}^-, \overline{x}^+]$, then there are at least two cutoff effort levels between which the players are indifferent.

7 Conclusion

In a dynamic game with multiple players experimenting on an unknown threshold, the features of the dynamics depend on the extent of information externality, the severity of breakdown, and patience. In the team problem, the time path of effort collapses to a very fast decline until it reaches some cutoff level. In contrast, for games with multiple players, the time path is gradual and smooth, asymptotically settling down at the same cutoff.

Conditional on no breakdown, the long-run effort level depends on the size of lumpy cost and patience, while the speed of decline changes with the number of players. As time goes on, the conditional hazard rate of a breakdown is decreasing.

References

ABREU, DILIP, AND FARUK GUL (2000), "Bargaining and Reputation," *Econometrica*, 68(1), 85-117.

ADMATI, ANAT R., AND MOTTY PERRY (1991), "Joint Projects wihout Commitment," *The Review of Economic Studies*, 58(2), 259-276.

BOLTON, PATRICK, AND CHRISTOPHER HARRIS (1999), "Stretegic Experimentation," *Econometrica*, 67(2), 349-374.

BONATTI, ALESSANDRO, AND JOHANNES HÖRNER (2011), "Collaborating," *American Economic Review*, 101, 632-663.

BONATTI, ALESSANDRO, AND JOHANNES HÖRNER (2013), "Searching for the Right Threshold," working paper.

CALLANDER, STEVEN (2011), "Searching and Learning by Trial and Error," *American Economic Review*, 101, 2277-2308.

CRIPPS, MARTIN, GODFREY KELLER, AND SVEN RADY (2005), "Strategic Experimentation with Exponential Bandits," *Econometrica*, 71(1), 39-68.

FUDENBERG, DREW, AND JEAN TIROLE (1986), "A Theory of Exit in Oligology," *Econometrica*, 54(4), 943-960.

HELLWIG, MARTIN, WOLFGANG LEININGER (1988), "Markov Perfect Equilibrium in Games of Perfect Information," Working Paper.

LANCASTER, KELVIN (1973), "The Dynamic Inefficiency of Capitalism," *Journal of Political Economy*, 81, 1098-1109.

LEVHARI, DAVID, LEONARD J. MIRMAN (1980), "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution," *Bell Journal of Economics*, 322-344.

MASKIN, ERIC, JEAN TIROLE (1988), "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs," *Econometrica*, 56(3), 549-569.

MASKIN, ERIC, JEAN TIROLE (1988), "A Theory of Dynamic Oligopoly, II: Kinked Demand Curve, and Edgeworth Cycles," *Econometrica*, 56(3), 571-599.

MATTHEWS, STEVEN A. (2013), "Achievable Outcomes of Dynamic Contribution Games," *Theoretical Economics*, 8, 365-403.

MOSCARINI, GUISEPPE (2005), "Job Matching and the Wage Distribution," *Econometrica*, 73(2), 481-516.

ROB, RAFAEL (1991), "Learning and Capacity Expansion under Demand Uncertainty," *The Review of Economic Studies*, 58(4), 655-675.

ROTHSCHILD, MICHAEL (1974), "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185-202.

A Appendix

A.1 Proof of Proposition 2

Proof. We prove part (a) first. If an equilibrium requires $\mathcal{V}(x) = 0$ for some $x \in (x_1^*, x_0]$, then HJB implies that $W_i(x) = \frac{p-x}{r}$. Also, $W_i(x') \geqslant \frac{p-x'}{r}$ for all $x' < x$, so that $W_i'(x) \leqslant -\frac{1}{r}$. Plug these into the first order condition to see that

$$\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) - L \right) - W_i'(x) \geqslant \frac{F(x) - Lrf(x)}{rF(x)} > 0$$

violating optimality.

Now we turn to part (b). Suppose that there is a positive measured set $A \subset (0, x_I^*)$ such that $\mathcal{V}(x) > 0$ for all $x \in A$. For $x \in A$ and for every $i = 1, \ldots, I$, solve $W_i'(x)$ from the HJB (7) and we have

$$W_i'(x) = -L \frac{\nu_i(x)}{\mathcal{V}(x)} \frac{f(x)}{F(x)} - \left(W_i(x) - \frac{p-x}{r} \right) \left(\frac{f(x)}{F(x)} + \frac{r}{\mathcal{V}(x)} \right)$$

Since Player i always has the choice to take null action $v_i = 0$, we must require $W_i(x) \geqslant \frac{p-x}{r}$, then the above becomes

$$W_i'(x) \leqslant -L \frac{\nu_i(x)}{\mathcal{V}(x)} \frac{f(x)}{F(x)} < -\frac{1}{r} \frac{I\nu_i(x)}{\mathcal{V}(x)}$$

Adding up the above for all i, we have

$$\frac{d}{dx} \sum_{i=1}^{I} W_i(x) < -\frac{I}{r}$$

On the other hand, for $x \in (0, x_I^*)\backslash A$, we have $\mathcal{V}(x) = 0$ and $\sum_{i=1}^I W_i(x) = -\frac{I(p-x)}{r}$. So,

$$
\begin{aligned}
& \lim_{x\to 0} \sum_{i=1}^I W_i(x) \\
= \quad & \sum_{i=1}^I W_i(x_I^*) - \lim_{x\to 0} \int_x^{x_I^*} \frac{d}{dx} \sum_{i=1}^I W_i(s)ds \\
= \quad & \sum_{i=1}^I W_i(x_I^*) - \lim_{x\to 0} \left(\int_A \frac{d}{dx} \sum_{i=1}^I W_i(s)ds + \int_{(0,x_I^*)\backslash A} \frac{d}{dx} \sum_{i=1}^I W_i(s)ds \right) \\
> \quad & \frac{I(p - x_I^*)}{r} + \lim_{x\to 0} \frac{I(x_I^* - x)}{r} \\
= \quad & \frac{Ip}{r}
\end{aligned}
$$

However, this consists a contradiction since $\lim_{x\to 0} \sum_{i=1}^I W_i(x) = \frac{Ip}{r}$ by Sandwich Theorem. Hence the measure of A is zero and $W_i(x) = \frac{p-x}{r}$. Plugging back $W_i(x)$ to the FOC implies that A is empty. ■

A.2 Proof of Proposition 3

Proof. Part (a) is directly implied by Proposition 2. We now turn to part (b). For $x \in (0, x_1^*)$, suppose there is a positive-measured set $A \subset (0, x_1^*)$ such that $\nu(x) > 0$ for all $x \in A$. Then for $x \in A$, first order condition for Player i reads

$$\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) - L \right) - W_i'(x) \geqslant 0$$

so that

$$W_i'(x) \leqslant \frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) - L \right) < -\frac{1}{r}$$

On the other hand, for $x \in (0, x_1^*)\backslash A$, we have $\nu(x) = 0$ and $W_i(x) = -\frac{p-x}{r}$. Also,

$W_i(x') \geqslant \frac{p-x'}{r}$ for all $x' < x$, so that $W_i'(x) \leqslant -1/r$. Hence,

$$\begin{aligned}\lim_{x\to 0} W_i(x) &= W_i(x_1^*) - \lim_{x\to 0}\int_x^{x_1^*} W_i'(s)ds \\ &= W_i(x_1^*) - \lim_{x\to 0}\left(\int_A W_i'(s)ds + \int_{(0,x_1^*)\setminus A} W_i'(s)ds\right) \\ &> \frac{p-x_1^*}{r} + \lim_{x\to 0}\frac{x_1^*-x}{r} \\ &= \frac{p}{r}\end{aligned}$$

This contradiction implies that the measure of A is zero, and furthermore that A is empty. ■

A.3 Proof of Theorem 4

Proof. That the proposed expression (9) for $\nu(\cdot)$ consists a symmetric equilibrium is guaranteed by verification Theorem taking other players' strategies as fixed.

The uniqueness of symmetric pure strategy equilibrium is shown below by following Tarski's fixed point Theorem. Notice first that combining (7) and (8) we have

$$BR_i(\nu_{-i})\begin{cases} = \overline{v} & \text{if } W_i(x) > \frac{p-x}{r} + \nu_{-i}(x)\frac{f(x)}{F(x)}\frac{L}{r}, \\ \in [0,\overline{v}] & \text{if } W_i(x) = \frac{p-x}{r} + \nu_{-i}(x)\frac{f(x)}{F(x)}\frac{L}{r}, \\ = 0 & \text{if } W_i(x) < \frac{p-x}{r} + \nu_{-i}(x)\frac{f(x)}{F(x)}\frac{L}{r}\end{cases}. \tag{19}$$

For any Lipschitz continuous function W with $-\frac{1}{r} \leqslant W' \leqslant 0$, define functional $\psi_1(\cdot)$ by $(\psi_1(W))(x) \equiv \min\left\{(I-1)\overline{v}, \frac{(rW(x)-(p-x))F(x)}{Lf(x)}\right\}$, and define functional $\psi_2(\nu_{-i})$ as the value function of a player when the aggregate action of other players is ν_{-i}. From (19), we know that W is a value function of a player in a symmetric pure strategy equilibrium if and only if W is a fixed point of $\psi \equiv \psi_2 \circ \psi_1$.

Evidently, ψ_1 is non-decreasing in W. The following lemma shows that ψ_2 is non-decreasing.

Lemma 18 *Consider problem* (7). *If $\tilde{\nu}_{-i} \geqslant \nu_{-i}$ for all x, then the respective solutions satisfy $\tilde{W}_i(x) \geqslant W_i(x)$ for all x.*

Proof. First, we bound $W_i(x)$ from below by $U_1(x)$. If not, then $\exists x_1$ s.t. $W_i(x_1) < U_1(x)$

and $W_i'(x_1) < U_1'(x_1)$. Following similar logic of Keller, Rady and Cripps (2005), we have the following inequality:

$$\begin{aligned}
& \max_{v_i \in [0,\overline{v}]} (p - x_1) + \frac{v_i f(x_1)}{F(x_1)} \left(\frac{p - x_1}{r} - U_1(x_1) - L \right) - V U_1'(x_1) \\
= \; & rU_1(x_1) \\
> \; & rW_i(x_1) \\
= \; & \max_{v_i \in [0,\overline{v}]} (p - x_1) + \frac{v_i f(x_1)}{F(x_1)} \left(\frac{p - x_1}{r} - W_i(x_1) - L \right) - v_i W_i'(x_1) \\
& + \frac{\nu_{-i} f(x_1)}{F(x_1)} \left(\frac{p - x_1}{r} - W_i(x_1) \right) - \nu_{-i} W_i'(x_1) \\
\geqslant \; & \max_{v_i \in [0,\overline{v}]} (p - x_1) + \frac{v_i f(x_1)}{F(x_1)} \left(\frac{p - x_1}{r} - U_1(x_1) - L \right) - v_i U_1'(x_1) \\
& + \frac{\nu_{-i} f(x_1)}{F(x_1)} \left(\frac{p - x_1}{r} - U_1(x_1) \right) - \nu_{-i} U_1'(x_1) \\
\geqslant \; & \max_{v_i \in [0,\overline{v}]} (p - x_1) + \frac{v_i f(x_1)}{F(x_1)} \left(\frac{p - x_1}{r} - U_1(x_1) - L \right) - v_i U_1'(x_1)
\end{aligned}$$

a contradiction. The last inequality follows from the fact that $\frac{f(x_1)}{F(x_1)} \left(\frac{p-x_1}{r} - U_1(x_1) \right) - U_1'(x_1) \geqslant 0$.

Next, we need to show that $\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) \right) - W_i'(x) \geqslant 0$.

If $v_i > 0$, then the FOC of (7) must require that $\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) - L \right) - W_i'(x) \geqslant 0$, hence we have $\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) \right) - W_i'(x) > 0$. If $v_i = 0$ and $\nu_{-i} = 0$, then $W_i(x) = \frac{p-x}{r}$ and $W_i'(x) \leqslant -\frac{1}{r}$, hence $\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) \right) - W_i'(x) \geqslant \frac{1}{r} > 0$. If $v_i = 0$ and $\nu_{-i} > 0$, then $\frac{f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) \right) - W_i'(x) = \frac{r}{\nu_{-i}} \left(W_i(x) - \frac{p-x}{r} \right) \geqslant 0$. In sum, the required inequality is true.

Finally,

$$\begin{aligned}
& \tilde{W}_i(x) \\
= \; & \max_{v_i \in [0,\overline{v}]} (p - x) + \frac{v_i f(x)}{F(x)} \left(\frac{p - x}{r} - \tilde{W}_i(x) - L \right) - v_i \tilde{W}_i'(x) \\
& + \frac{\tilde{\nu}_{-i} f(x)}{F(x)} \left(\frac{p - x}{r} - \tilde{W}_i(x) \right) - \tilde{\nu}_{-i} \tilde{W}_i'(x) \\
\geqslant \; & \max_{v_i \in [0,\overline{v}]} (p - x) + \frac{v_i f(x)}{F(x)} \left(\frac{p - x}{r} - \tilde{W}_i(x) - L \right) - v_i \tilde{W}_i'(x) \\
& + \frac{\nu_{-i} f(x)}{F(x)} \left(\frac{p - x}{r} - \tilde{W}_i(x) \right) - \nu_{-i} \tilde{W}_i'(x)
\end{aligned}$$

Since $\tilde{W}_i(0) = W_i(0) = (p-x)/r$, if ever $\tilde{W}_i < W_i$, there must exist x_1 s.t. $\tilde{W}_i(x_1) < W_i(x_1)$ and $\tilde{W}_i'(x_1) < W_i'(x_1)$. Then

$$\begin{aligned}
& \tilde{W}_i(x_1) \\
\geqslant \ & \max_{v_i \in [0,\overline{v}]} (p-x) + \frac{v_i f(x)}{F(x)} \left(\frac{p-x}{r} - \tilde{W}_i(x) - L \right) - v_i \tilde{W}_i'(x) \\
& + \frac{\nu_{-i} f(x)}{F(x)} \left(\frac{p-x}{r} - \tilde{W}_i(x) \right) - \nu_{-i} \tilde{W}_i'(x) \\
\geqslant \ & \max_{v_i \in [0,\overline{v}]} (p-x) + \frac{v_i f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) - L \right) - v_i W_i'(x) \\
& + \frac{\nu_{-i} f(x)}{F(x)} \left(\frac{p-x}{r} - W_i(x) \right) - \nu_{-i} W_i'(x) \\
= \ & W_i(x_1)
\end{aligned}$$

a contradiction. ■

Hence, it follows from Tarski's fixed point Theorem that $\psi = \psi_2 \circ \psi_1$ has minimal and maximal fixed points W_- and W_+. For W_-, we know from Proposition 3 that

$$W_-'(x) = \begin{cases} -\frac{1}{r} & \text{if } x \in [0, x_1^*] \\ \min\{\frac{p-x-rW_-(x)}{I\overline{v}} + \frac{(I-1)Lf(x)}{IF(x)}, 0\} - \frac{(-p+Lr+x+rW_-(x))f(x)}{rF(x)} & \text{if } x \in [x_1^*, x_0] \end{cases}$$

Similarly, there exists a unique cutoff $x_+ \in [0, x_1^*]$ for W_+. Hence the difference $\overline{W} \equiv W_+ - W_-$ satisfies

$$\overline{W}'(x) = \begin{cases} 0 & \text{if } x \in [0, x_1^*] \\ -\frac{\overline{W}(x)f(x)}{F(x)} + \min\{\frac{p-x-rW_+(x)}{I\overline{v}} + \frac{(I-1)Lf(x)}{IF(x)}, 0\} & \\ \quad - \min\{\frac{p-x-rW_-(x)}{I\overline{v}} + \frac{(I-1)Lf(x)}{IF(x)}, 0\} & \text{if } x \in (x_1^*, x_0] \end{cases}.$$

Therefore, $\overline{W}(x) \geqslant 0$ and $\overline{W}'(x) \leqslant 0$ for all $x \in [0, x_0]$. Meanwhile, $\overline{W}(0) = 0$ because $W_+(0) = W_-(0) = \frac{p}{r}$. So $\overline{W}(x) = 0$ for all $x \in [0, x_0]$, implying uniqueness. ■

A.4 Proof of Proposition 6

Proof. Since $\lim_{x \downarrow x_1^*} \frac{d}{dx} \left(\frac{F(x)}{f(x)} \right) = b$, we have $\lim_{x \downarrow x_1^*} \nu(x) = \frac{b(x-x_1^*)^2}{2L(I-1)}$. For any $A > 1$, there exists a $\hat{x}$ s.t. $\frac{1}{A} \frac{b(x-x_1^*)^2}{2L(I-1)} < \nu(x) < A \frac{b(x-x_1^*)^2}{2L(I-1)}$ for $x \in [x_1^*, \hat{x}]$. Denote $\hat{t}$ as the solution to $\hat{x} = x(\hat{t})$. Starting from the point $(\hat{t}, \hat{x})$, the path $x(t)$ is sandwiched by paths where $\nu(x)$ is replaced by $\frac{1}{A}\frac{b(x-x_1^*)^2}{2L(I-1)}$ and $A\frac{b(x-x_1^*)^2}{2L(I-1)}$, respectively. Hence, we have

$$\frac{2(I-1)Lt(\hat{x}-x_1^*)}{2(I-1)L+AbI(t-\hat{t})(\hat{x}-x_1^*)} < \frac{x(t)-x_1^*}{1/t} < \frac{2(I-1)Lt(\hat{x}-x_1^*)}{2(I-1)L+A^{-1}bI(t-\hat{t})(\hat{x}-x_1^*)} \tag{20}$$

for $t > \hat{t}$, so that

$$\limsup_{t\to\infty} \frac{x(t)-x_1^*}{1/t} \leqslant \frac{2(I-1)LA}{bI}$$
$$\liminf_{t\to\infty} \frac{x(t)-x_1^*}{1/t} \geqslant \frac{2(I-1)L}{AbI}$$

Since the above hold for any $A > 1$, we know $\lim_{t\to\infty} \frac{x(t)-x_1^*}{1/t} = \frac{2(I-1)L}{bI}$. ■

A.5 Proof of Lemma 8

Proof. Notice that x_1^* is a function of L and r. Taking derivative of $\nu(x)$ w.r.t. L and r gives

$$\frac{\partial\nu(x)}{\partial L} = -\frac{\int_{x_1^*}^{x} F(s)ds}{(I-1)L^2 f(x)} - \frac{\partial x_1^*}{\partial L}\frac{F(x_1^*)-Lrf(x_1^*)}{(I-1)Lf(x)} = -\frac{\int_{x_1^*}^{x} F(s)ds}{(I-1)L^2 f(x)} < 0$$
$$\frac{\partial\nu(x)}{\partial r} = -\frac{F(x)-F(x_1^*)}{(I-1)f(x)} - \frac{\partial x_1^*}{\partial r}\frac{F(x_1^*)-Lrf(x_1^*)}{(I-1)Lf(x)} = -\frac{F(x)-F(x_1^*)}{(I-1)f(x)} < 0$$

Since both $\frac{1}{I-1}$ and $\frac{I}{I-1}$ are decreasing in I, the statement for I is obvious. ■

A.6 Proof of Proposition 10

Proof. The key step of the proof is Lemma 19 below. In order to understand the structure of MPE's in Proposition 10, consider the following system of difference equations with generic variables y_k and z_k.

$$(p-y_k)\tilde{\Delta}+\delta z_k = (p-y_{k+1})\tilde{\Delta} - L\frac{F(y_k)-F(y_{k+1})}{F(y_k)} + \delta\left[z_{k+1}\frac{F(y_{k+1})}{F(y_k)} + \frac{p-y_k}{r}\frac{F(y_k)-F(y_{k+1})}{F(y_k)}\right], \tag{21}$$

$$z_k = (p-y_k)\tilde{\Delta} + \delta\left[[(p-y_{k+1})\tilde{\Delta}+\delta z_{k+1}]\frac{F(y_{k+1})}{F(y_k)} + \frac{p-y_k}{r}\frac{F(y_k)-F(y_{k+1})}{F(y_k)}\right], \tag{22}$$

$$y_0 \in [x_1^*, x_0] \text{ is given.} \tag{23}$$

One can think of $\{y_k\}_{k=0}^{\infty}$ as a sequence of critical effort levels and $\{z_k\}_{k=0}^{\infty}$ as the sequence of corresponding payoffs when faced with the critical effort levels and the player is currently *not* the mover. The MPE we look for has the skimming property such that starting from a state $x \in (y_{k+1}, y_k]$, the current mover takes action $\frac{x-y_{k+1}}{\Delta}$ to bring down the state to the highest critical level strictly below x, namely y_{k+1}. Equation (21) is the indifference condition that facing state $x = y_k$, the player is indifferent between staying at y_k and moving one step down to y_{k+1}, given the continuation play prescribed by the MPE. Equation (22) is the promise keeping condition saying that the payoff of the non-mover facing state y_k is the weighted sum of current payoff and continuation payoff, where the indifference condition is already embodied in the continuation payoff. Now we state Lemma 19.

Lemma 19 *Suppose Δ is small. There exists a unique $z_0 \geqslant 0$ such that the solution to the difference equation system* (21)-(23) *has the property that y_k monotonically decreases and* $\lim_{k\to\infty} y_k = x_1^*$.

Proof.

Step 1: Change of variables: $z_k = \frac{u_k}{F(y_k)} + \frac{p-y_k}{r}$.

Noting also that $\tilde{\Delta} = (1-\delta)/r$, the system is simplified to

$$u_{k+1} = \frac{ru_k - \delta(y_k - y_{k+1})F(y_{k+1})}{\delta^2 r} \tag{24}$$

$$\delta[Lr - (1-\delta)(y_k - y_{k+1})][F(y_k) - F(y_{k+1})] = (1-\delta^2)ru_k \tag{25}$$

Now, (24), (25) and (23) consist a new difference equation system with generic variables y_k and u_k. For a fixed y_0, there is one-to-one mapping from u_0 to z_0, so we want to find the unique u_0 such that y_k monotonically converges to x_1^*. We can immediately rule out the case $u_0 < 0$. To see why, note that $y_0 - y_1 < x_0 - x_1^*$ is bounded, so for δ close enough to 1 (Δ small), $Lr - (1-\delta)(y_0 - y_1) > 0$, and hence from (25) we have $u_0 \geqslant 0$.

Step 2: Induction on the new system.

First we examine conditions for (y_k, u_k) s.t. (y_{k+1}, u_{k+1}) exists as the unique solution to (24) and (25). From (25) define

$$J(y_k, y_{k+1}, u_k) \equiv \delta[Lr - (1-\delta)(y_k - y_{k+1})][F(y_k) - F(y_{k+1})] - (1-\delta^2)ru_k$$

so that

$$\frac{\partial J}{\partial y_{k+1}} = (1-\delta)[F(y_k) - F(y_{k+1})] - [Lr - (1-\delta)(y_k - y_{k+1})]f(y_{k+1})$$
$$\frac{\partial J}{\partial y_k} = -(1-\delta)[F(y_k) - F(y_{k+1})] + [Lr - (1-\delta)(y_k - y_{k+1})]f(y_k)$$
$$\frac{\partial J}{\partial u_k} = -(1-\delta^2)r$$

Now, $J(y_k, y_k, u_k) \leqslant 0$ and moreover, $\frac{\partial J}{\partial y_{k+1}} < 0$ for $y_{k+1} \in [x_1^*, y_k]$ when $y_k > x_1^*$ and δ is close enough to 1 (remember that f is bounded below by $\underline{f} > 0$). By Intermediate Value Theorem, there exists a unique $y_{k+1} \in [x_1^*, y_k]$ if and only if $J(y_k, x_1^*, u_k) \geqslant 0$, i.e. u_k is not too large given y_k. Having pinned down y_{k+1}, we immediately determine u_{k+1} by (24).

The following lines show that $u_k \geqslant 0 \Rightarrow u_{k+1} \geqslant 0$ if $y_{k+1} \in [x_1^*, y_k]$ exists.

$$\begin{aligned}
\delta^2 r u_{k+1} &= r u_k - \delta(y_k - y_{k+1})F(y_{k+1}) \geqslant r u_k - \frac{\delta}{\underline{f}}[F(y_k) - F(y_{k+1})]F(y_{k+1}) \\
&= r u_k \left(1 - \frac{(1-\delta^2)F(y_{k+1})}{\underline{f}[Lr - (1-\delta)(y_k - y_{k+1})]}\right) \\
&\geqslant r u_k \left(1 - \frac{1-\delta^2}{\underline{f}[Lr - (1-\delta)x_0]}\right) \geqslant 0
\end{aligned}$$

where the last inequality follows when δ is close to 1.

With the above observations, we define the transition function from tuple to tuple:

$$\Gamma(y_k, u_k) \equiv (\Gamma_y(y_k, u_k), \Gamma_u(y_k, u_k)) \equiv (y_{k+1}, u_{k+1})$$

where (y_{k+1}, u_{k+1}) is the solution to (24) (25) satisfying $y_{k+1} \in [x_1^*, y_k]$, if it exists. Otherwise, the function returns some arbitrary vector, say $(-1, -1)$, to indicate nonexistence. $\Gamma^{(n)} \equiv (\Gamma_y^{(n)}(y_k, u_k), \Gamma_u^{(n)}(y_k, u_k))$ is the function Γ applied n times.

Step 3: Uniform upper bound on u_0 such that $\Gamma_y^{(\infty)}(y_0, u_0) \geqslant x_1^*$.

Fix (y_0, u_0), we want to bound the locus of $\Gamma^n(y_0, u_0)$ from below on the $y - u$ plane. Define a lower bound function:

$$G_L(y) \equiv u_0 - \frac{1}{\delta r}\int_y^{y_0} [F(s) - Lrf(s)]ds \tag{26}$$

where $y_0 \geqslant x_1^*$, with $G_L(y_0) = u_0$. If we can show that $u_k \geqslant G_L(y_k) \Rightarrow \Gamma_u(y_k, u_k) \geqslant$

$G_L(\Gamma_y(y_k,u_k))$ provided $\Gamma_y(y_k,u_k)\in[x_1^*,y_k]$, then by induction $\Gamma_u^{(n)}(y_k,u_k)\geqslant G_L(\Gamma_y^{(n)}(y_k,u_k))$ for all n, provided $\Gamma_y^{(n)}(y_k,u_k)\in[x_1^*,\Gamma_y^{(n-1)}(y_k,u_k)]$. Let $u_k=G_L(y_k)+\varepsilon$ where $\varepsilon\geqslant 0$. Denote $y_{k+1}=\Gamma_y(y_k,u_k)$ and $u_{k+1}=\Gamma_u(y_k,u_k)=\frac{ru_k-\delta(y_k-y_{k+1})F(y_{k+1})}{\delta^2 r}$ (by (24)). The claim above is true if

$$\begin{aligned}
&\frac{r[G_L(y_k)+\varepsilon]-\delta(y_k-y_{k+1})F(y_{k+1})}{\delta^2 r}\geqslant G_L(y_{k+1})=G_L(y_k)-\frac{1}{\delta r}\int_{y_{k+1}}^{y_k}[F(s)-Lrf(s)]ds\\
\Leftrightarrow\quad &(1-\delta^2)G_L(y_k)\geqslant-\frac{\delta}{r}\int_{y_{k+1}}^{y_k}[F(s)-F(y_{k+1})-Lrf(s)]ds-\varepsilon \qquad (27)
\end{aligned}$$

On the other hand, (25) gives

$$G_L(y_k)=u_k-\varepsilon=\frac{\delta}{(1-\delta^2)r}[Lr-(1-\delta)(y_k-y_{k+1})][F(y_k)-F(y_{k+1})]-\varepsilon$$

so this together with (27) yields

$$(1-\delta)(y_k-y_{k+1})[F(y_k)-F(y_{k+1})]-\int_{y_{k+1}}^{y_k}[F(s)-F(y_{k+1})]ds\leqslant 0 \qquad (28)$$

because it holds for all $\varepsilon\geqslant 0$.

Let $f_m\equiv\min_{x\in[y_{k+1},y_k]}f(x)\geqslant\underline{f}$, then $F(s)-F(y_{k+1})\geqslant f_m(s-y_{k+1})$. By Lipschitz continuity of f, we have $\frac{F(y_k)-F(y_{k+1})}{y_k-y_{k+1}}\leqslant f_m+\kappa(y_k-y_{k+1})\leqslant f_m+\kappa x_0$. So, a sufficient condition for (28) is

$$\begin{aligned}
&(1-\delta)(y_k-y_{k+1})^2(f_m+\kappa x_0)-f_m\int_{y_{k+1}}^{y}(s-y_{k+1})ds\leqslant 0\\
\Leftrightarrow\quad &\delta\geqslant 1-\frac{f_m}{2(f_m+\kappa x_0)}
\end{aligned}$$

Since $f_m\geqslant\underline{f}$, a uniform sufficient condition is $\delta\geqslant 1-\frac{\underline{f}}{2(\underline{f}+\kappa x_0)}$.

Note that $G_L(\cdot)$ is strictly increasing in $[x_1^*,y_0]$. Also, $G_L(x_1^*)=u_0-\frac{1}{\delta r}\int_{x_1^*}^{y_0}F(s)ds+\frac{L}{\delta}[F(y_0)-F(x_1^*)]>0$ if $u_0>\overline{u}\equiv\frac{y_0-x_1^*}{\delta r}$.

Hence, if u_0 is too big for a given y_0, then $u_k>G_L(y_k)\geqslant G_L(x_1^*)>0$ whenever $y_k\in[x_1^*,x_0]$. However, in order to have $y_\infty=x_1^*$, we must have $u_\infty=0$ by (25), a contradiction. Therefore, given any $y_0\in[x_1^*,x_0]$, u_0 has a uniform upper bound independent of δ. Moreover, because $G_L'(y)>0$ for all $y\in[x_1^*,y_0]$, we can restrict attention to points $(y,u)\in[x_1^*,y_0]\times[0,\overline{u}]$ for further analysis.

Step 4: Uniform lower bound on u_0 such that $\Gamma_y^{(\infty)}(y_0, u_0) \leqslant x_1^*$.

Fix (y_0, u_0), we want to bound the locus of $\Gamma^n(y_0, u_0)$ from above on the $y - u$ plane. To achieve this, define

$$G_H(y) = u_0 + a(1-\delta)(y_0 - y) - \frac{1}{\delta r}\int_y^{y_0} [F(s) - Lrf(s)]ds \tag{29}$$

where $y_0 \geqslant x_1^*$, a is a positive constant to be determined later, and $G_H(y_0) = u_0$. Because of induction, we only need to show that $u \leqslant G_H(y) \Rightarrow \Gamma_u(y, u) \leqslant G_H(\Gamma_y(y, u))$ if $\Gamma_y(y, u) \in [x_1^*, y]$ and $u \in [0, \overline{u}]$. I claim that this is true for some $a > 0$. Let $u_k = G_H(y_k) - \varepsilon$ where $\varepsilon \geqslant 0$. Denote $y_{k+1} = \Gamma_y(y_k, u_k)$ and $u_{k+1} = \Gamma_u(y_k, u_k) = \frac{ru_k - \delta(y_k - y_{k+1})F(y_{k+1})}{\delta^2 r}$ as before. The claim is true if

$$\begin{aligned} & \frac{r[G_H(y_k) - \varepsilon] - \delta(y_k - y_{k+1})F(y_{k+1})}{\delta^2 r} \\ \leqslant \quad & G_H(y_{k+1}) = G_H(y_k) + a(1-\delta)(y_k - y_{k+1}) - \frac{1}{\delta r}\int_{y_{k+1}}^{y_k} [F(s) - Lrf(s)]ds \end{aligned}$$

which is true if the following sufficient condition is satisfied:

$$(1-\delta^2)G_H(y_k) \leqslant a\delta^2(1-\delta)(y_k - y_{k+1}) - \frac{\delta}{r}\int_{y_{k+1}}^{y_k} [F(y_k) - F(y_{k+1}) - Lrf(s)]ds + \varepsilon \tag{30}$$

On the other hand, (25) gives

$$G_H(y_k) = u_k + \varepsilon = \frac{\delta}{(1-\delta^2)r}[Lr - (1-\delta)(y_k - y_{k+1})][F(y_k) - F(y_{k+1})] + \varepsilon$$

so that (30) reduces to

$$\begin{aligned} & a \geqslant \frac{[F(y_k) - F(y_{k+1})](y_k - y_{k+1}) - r\varepsilon}{r(1-\delta)(y_k - y_{k+1})} \text{ for all } \varepsilon \geqslant 0 \\ \Rightarrow \quad & a \geqslant \frac{(1+\delta)u_k}{\delta[Lr - (1-\delta)(y_k - y_{k+1})]} \end{aligned}$$

For δ close to 1, it is sufficient to set $a = \frac{4u_0}{Lr}$. With this value of a, $u_0 = \max_{y \in [x_1^*, y_0]}$, and G_H is indeed an upper bound for the sequence $\Gamma^{(n)}(y_0, u_0)$ when δ is close to 1.

Note that

$$\begin{aligned} G_H(x_1^*) &= u_0 + a(1-\delta)(y_0 - x_1^*) - \frac{1}{\delta r}\int_{x_1^*}^{y_0} f(s)\left(\frac{F(s)}{f(s)} - \frac{F(x_1^*)}{f(x_1^*)}\right) ds \\ &\leqslant u_0 + a(1-\delta)(y_0 - x_1^*) - \frac{1}{\delta r}\int_{x_1^*}^{y_0} \underline{f}(s - x_1^*)\underline{b} ds \\ &= u_0 + \frac{4u_0}{Lr}(1-\delta)(y_0 - x_1^*) - \frac{(y_0 - x_1^*)^2 \underline{f}\underline{b}}{2\delta r} \end{aligned}$$

so that when δ is close to 1, $G_H(x_1^*) < 0$ if $u_0 < \frac{(y_0 - x_1^*)^2 \underline{f}\underline{b}}{4r}$. In order to have $y_\infty = x_1^*$, we require $u_\infty = 0$, but the point $(x_1^*, 0)$ is above the graph of G_H, a contradiction. Actually, if $u_0 < \frac{(y_0 - x_1^*)^2 \underline{f}\underline{b}}{4r}$, we know that for δ close to 1, $J(y_0, x_1^*, u_0) > 0$, so that $y_1 \in (x_1^*, y_0]$. By induction, $y_k \in (x_1^*, y_0]$ for any k, meaning that $\{y_k\}_{k=0}^\infty$ is a decreasing sequence with lower bound x_1^*, admitting a limit $y_\infty \in (x_1^*, y_0]$. Therefore, fix any $y_0 \in [x_1^*, x_0]$, a too small u_0 leads to $y_\infty > x_1^*$.

Step 5: Order preserving property of $\Gamma(\cdot, \cdot)$.

Order preserving property means that for any two different points $(y, u), (y', u') \in [x_1^*, x_0] \times [0, \overline{u}]$ with $y' \leqslant y$ and $u' \geqslant u$, $\Gamma_y(y', u') < \Gamma_y(y, u)$ and $\Gamma_u(y', u') > \Gamma_u(y, u)$, provided $\Gamma_y(y', u'), \Gamma_y(y, u) \geqslant x_1^*$.

We want to prove the above for $u, u' <$. Given existence, recall from Step 2 that $\frac{\partial J}{\partial y_{k+1}} < 0$, $\frac{\partial J}{\partial y_k} > 0$ and $\frac{\partial J}{\partial u_k} < 0$ for δ close to 1. By Implicit Function Theorem, $\frac{\partial y_{k+1}}{\partial y_k} > 0$ and $\frac{\partial y_{k+1}}{\partial u_k} < 0$.

Moreover, from (24) we use the chain rule to get

$$\frac{\partial u_{k+1}}{\partial u_k} = \frac{1}{\delta^2} - \frac{1}{\delta r}[(y_k - y_{k+1})f(y_{k+1}) - F(y_{k+1})]\frac{\partial y_{k+1}}{\partial u_k} \tag{31}$$

Note that while $(y_k - y_{k+1})f(y_{k+1}) - F(y_{k+1})$ is bounded, $\frac{\partial y_{k+1}}{\partial u_k}$ uniformly converges to zero as $\delta \to 1$, so $\frac{\partial y_{k+1}}{\partial u_k} > 0$ for δ close to 1.

Finally,

$$\frac{\partial u_{k+1}}{\partial y_k} = -\frac{F(y_{k+1})}{\delta r} - \frac{1}{\delta r}[(y_k - y_{k+1})f(y_{k+1}) - F(y_{k+1})]\frac{\partial y_{k+1}}{\partial y_k}$$

If $\frac{\partial y_{k+1}}{\partial y_k} \leqslant 1$, then $\frac{\partial u_{k+1}}{\partial y_k} < 0$. Otherwise, $\frac{\partial y_{k+1}}{\partial y_k} < \frac{Lrf(y_k) - (1-\delta)}{Lrf(y_{k+1}) - (1-\delta)}$ when δ is close to 1. By the assumption on the hazard rate, we have

$$\begin{aligned} & \frac{F(y_k)}{f(y_k)} > \frac{F(y_{k+1})}{f(y_{k+1})} + \underline{b}(y_k - y_{k+1}) \\ \Rightarrow \quad & F(y_{k+1})[f(y_k) - f(y_{k+1})] < f(y_{k+1})[F(y_k) - F(y_{k+1})] - \underline{b}(y_k - y_{k+1}) f(y_k) f(y_{k+1}) \end{aligned}$$

so that

$$\begin{aligned} \frac{\partial u_{k+1}}{\partial y_k} &< \frac{f(y_{k+1})(y_k - y_{k+1})}{\delta r[f(y_{k+1})Lr - (1-\delta)]} \left[\left(\frac{F(y_k) - F(y_{k+1})}{y_k - y_{k+1}} - f(y_k) - \underline{b} f(y_k) \right) Lr + (1-\delta) \right] \\ &< \frac{f(y_{k+1})(y_k - y_{k+1})}{\delta r[f(y_{k+1})Lr - (1-\delta)]} \left[\left(\kappa(y_k - y_{k+1}) - \underline{b} f(y_k) \right) Lr + (1-\delta) \right] \\ &< \frac{f(y_{k+1})(y_k - y_{k+1})}{\delta r[f(y_{k+1})Lr - (1-\delta)]} \left[\left(\frac{\kappa r \overline{u}(1-\delta^2)}{\delta[Lr - (1-\delta)(y_k - y_{k+1})]} - \underline{b} f(y_k) \right) Lr + (1-\delta) \right] \\ &< -\frac{\underline{b} f(y_k - y_{k+1})}{2r} \end{aligned}$$

for δ close to 1.

With the signs of the four partial derivatives, we have $\Gamma_y(y', u') < \Gamma_y(y, u)$ and $\Gamma_u(y', u') > \Gamma_u(y, u)$ if $u' \geqslant u$, $y' \leqslant y$ and $(y', u') \neq (y, u)$. Iteration forward gives us the desired ordering for the whole sequence.

Step 6: For those u_0 s.t. $y_\infty \in [x_1^*, y_0]$, y_∞ is strictly decreasing in u_0.

Suppose this is not true, then there exist two initial points (y_0, u_0) and (y_0, u_0') with $u_0' > u_0$ but $y_\infty' = y_\infty \geqslant x_1^*$ (by Step 5, $y_\infty' > y_\infty$ is impossible). We will show that this cannot happen.

The idea is that if $y_\infty' = y_\infty$ then $y_k - y_k'$ will be small for any k, contradicting the initial difference $F(y_1) - F(y_1') > 0$. Formally, for any $s \geqslant 0$, by (25)

$$\begin{aligned} & F(y_s') - F(y_{s+1}') \\ = \quad & \frac{Lr - (1-\delta)(y_s - y_{s+1})}{Lr - (1-\delta)(y_s' - y_{s+1}')} [F(y_s) - F(y_{s+1})] \frac{u_s'}{u_s} \\ > \quad & \frac{Lr - (1-\delta)(y_s - y_{s+1})}{Lr - (1-\delta)(y_s' - y_{s+1}')} [F(y_s) - F(y_{s+1})] \\ > \quad & [F(y_s) - F(y_{s+1})] \left(1 - \frac{(1-\delta)(y_s - y_s')}{Lr - (1-\delta)x_0} \right) \end{aligned}$$

using the fact that $u_s' > u_s$ and $y_s > y_s'$.

For an arbitrary $k \geqslant 1$, adding up the above inequality on the far left and far right sides for s from k to ∞, we have

$$
\begin{aligned}
& F(y_k') - F(y^*) \\
> \;& F(y_k) - F(y^*) - \frac{1-\delta}{Lr-(1-\delta)x_0}\sum_{s=k}^{\infty}[F(y_s)-F(y_{s+1})](y_s - y_s') \\
> \;& F(y_k) - F(y^*) - \frac{1-\delta}{Lr-(1-\delta)x_0}x_0 \qquad (32)
\end{aligned}
$$

which means

$$
\begin{aligned}
& F(y_k) - F(y_k') < \frac{1-\delta}{Lr-(1-\delta)x_0}x_0 \\
\Rightarrow \;& y_k - y_k' < Ax_0 \qquad (33)
\end{aligned}
$$

where $A \equiv \frac{1-\delta}{\underline{f}[Lr-(1-\delta)x_0]}$.

Note that in deriving (32) we use the initial bound that $y_s - y_s' < x_0$ for all s, and arrive at another bound that that $y_k - y_k' < Ax_0$ for all k. One can iterate between inequalities (32) and (33) for arbitrarily many (n) rounds to get $y_k - y_k' < A^n x_0$ for all k. The factor A is smaller than 1 if δ is close to 1, resulting in $y_k - y_k' = 0$. But we know from Step 5 that for $u_0' > u_0$, it must be the case that $y_k > y_k'$. This contradiction means that $y_\infty' < y_\infty$.

Step 7: There is a unique u_0 such that $\Gamma_y^{(\infty)}(y_0, u_0) = x_1^*$.

For any given $y_0 \in [x_1^*, x_0]$, define $\mathcal{U} \equiv \{u_0 \in [0, \overline{u}] : \Gamma_y^{(\infty)}(y_0, u_0) \in [x_1^*, y_0]\}$. From Steps 3, 4 and 5 we know that this set is non-empty, bounded above, and has the property $u \in \mathcal{U} \Rightarrow u' \in \mathcal{U}$ for all $u' \in [0, u]$. Hence, $\mathcal{U}$ is an interval $[0, u_0^*]$ or $[0, u_0^*)$ for some $u_0^* > 0$.

It can be shown that $u_0^* = \sup \mathcal{U} \in \mathcal{U}$. To see this, let $u_0 > u_0^*$ so that $\Gamma_y^{(\infty)}(y_0, u_0) < x_1^*$, and hence there exists a smallest $k \geqslant 1$ s.t. $\Gamma_y^{(k)}(y_0, u_0) < x_1^*$. This means $J(y_{k-1}, x_1^*, u_{k-1}) < 0$. Notice that fix y_0, $(y_{k-1}, u_{k-1}) = \Gamma^{(k-1)}(y_0, u_0)$ is a continuous function of u_0, so if $u_0' < u_0$ is close enough to u_0, we still have $J(y_{k-1}', x_1^*, u_{k-1}') < 0$. Therefore, $\mathcal{U}$ is closed and $u_0^* \in \mathcal{U}$.

It remains to show that $\Gamma_y^{(\infty)}(y_0, u_0^*) = x_1^*$. Suppose towards contradiction that $\Gamma_y^{(\infty)}(y_0, u_0^*) = \hat{x} > x_1^*$. Let $(y_k, u_k) = \Gamma^{(k)}(y_0, u_0^*)$. Fix an arbitrary $\eta > 0$. Because of the convergence, there is a K such that $|y_K - \hat{x}| < \frac{\eta}{2}$ and $0 < u_K < \frac{\eta}{2}$. On the other hand, because of the continuity of $\Gamma^{(K)}(\cdot, \cdot)$, we know that there is a $\varepsilon > 0$ small enough s.t. $|y_K' - y_K| < \frac{\eta}{2}$ and $|u_K' - u_K| < \frac{\eta}{2}$, where $(y_K', u_K') = \Gamma^{(K)}(y_0, u_0^* + \varepsilon)$. Hence, $|y_K' - \hat{x}| < \eta$ and $0 < u_K' < \eta$. By virtue of Step 4, if $u_K' < \frac{(y_K' - x_1^*)^2 \underline{f} b}{4r}$ then $\Gamma^{(\infty)}(y_{k_1}', u_{k_1}')$ exists. This condition is satisfied when the arbitrarily fixed η is small enough. Step 4 then implies $u_0^* + \varepsilon \in \mathcal{U}$, contradicting the definition of u_0^*. Therefore, $\Gamma_y^{(\infty)}(y_0, u_0^*) = x_1^*$. By Step 6, it is the unique u_0 s.t. $y_\infty = x_1^*$. ■

From above we know that every $y_0 \in [x_1^*, x_0]$ pins down a unique $u_0^*(y_0)$. In the following, we aim to show that for any $y_0 \in [\underline{y}_0, x_0]$, the strategy defined in Proposition 10 with $y_k = \Gamma_y^{(k)}(y_0, u_0^*(y_0))$ consists an MPE, where $\underline{y}_0 \equiv \Gamma_y(x_0, u_0^*(x_0))$. For any $x \in (x_1^*, x_0]$, we must have $x \in (y_k, y_{k-1}]$ for some $k \geqslant 0$.

Step 1: Verify that for every $k \geqslant 0$, the payoff of the current mover facing state $x \in (y_k, y_{k-1}]$ and brings new state x' is decreasing in x' on every interval (y_{k+s}, y_{k+s-1}) for all $s \geqslant 1$ and on interval (y_k, x), where $x' = x - v\Delta$ and $y_{-1} = x_0$.

The payoff of the current mover if she brings new state $x' \in (y_{k+s}, y_{k+s-1})$ is

$$\begin{aligned} U(x';x) &= (p-x')\tilde{\Delta} - L\frac{F(x)-F(x')}{F(x)} \\ &\quad +\delta\left[\frac{F(x)-F(x')}{F(x)}\frac{p-x}{r} + \frac{F(x')}{F(x)}w(x')\right] \end{aligned}$$

where

$$\begin{aligned} w(x') &= (p-x')\tilde{\Delta} \\ &\quad +\delta\left[\frac{F(x')-F(y_{k+s})}{F(x')}\frac{p-x'}{r} + \frac{F(y_{k+s})}{F(x')}[(p-y_{k+s})\tilde{\Delta}+\delta z_{k+s}]\right] \end{aligned}$$

and $z_{k+s} = \frac{u_{k+s}}{F(y_{k+s})} + \frac{p-y_{k+s}}{r}$, $(y_{k+s}, u_{k+s}) = \Gamma^{(k+s)}(y_0, u_0^*)$.

Plugging in $\tilde{\Delta} = (1-\delta)/r$, we have

$$\begin{aligned} & \frac{dU(x';x)}{dx} > 0 \\ \Leftrightarrow\quad & -\delta F(x') + \delta^2 F(y_{k+s}) + (Lr + \delta(x-x'))f(x) > 0 \end{aligned}$$

which is true when δ is close enough to 1. So, for $x \in (y_{k+s}, y_{k+s-1}]$, $U(x';x)$ is increasing in x'. The above analysis also holds for $x' \in (y_k, x)$. In particular, $U(x';x) \leqslant U(x;x)$ when $x' \in (y_k, x]$.

Step 2: Verify that the current mover facing state $x \in (y_k, y_{k-1}]$ prefers y_k to x.

The equilibrium action leading to y_k yields

$$\begin{aligned} U(y_k;x) &= (p-y_k)\tilde{\Delta} - L\frac{F(x)-F(y_k)}{F(x)} \\ &\quad +\delta\left[\frac{F(x)-F(y_k)}{F(x)}\frac{p-M}{r} + \frac{F(y_k)}{F(x)}z_k\right] \end{aligned}$$

so

$$\begin{aligned} & U(y_k;x) \geqslant U(x;x) \\ \Leftrightarrow\quad & -(1-\delta^2)\delta r u_k + [Lr-(1-\delta)(x-y_k)]F(x) - [Lr+(1-\delta)\delta(x-y_k)]F(y_k) \leqslant 0 \quad (34) \end{aligned}$$

where $u_k = (z_k - (p-y_k)/r)[F(y_k) - F(m)]$. Recall from (24) and (25) that (34) holds with equality when $x = y_{k-1}$ (provided $k \geqslant 1$). Taking the derivative of the left hand side of (34) w.r.t. x gives

$$-(1-\delta)[F(x)+\delta F(y_k)] + [Lr-(1-\delta)(x-y_k)]f(x) > 0$$

when δ is close to 1. That means, the left hand side of (34) is non-positive for $x \in (y_k, y_{k-1}]$, i.e. $U(y_k;x) \geqslant U(x;x)$ when $k \geqslant 1$.

The only caveat is for the case $k = 0$. Because x_0 does not belong to the sequence $\{y_k\}_{k=0}^{\infty}$, the indifference condition does not hold at $x = x_0$. In order to have $U(y_0;x) \geqslant U(x;x)$ when $x \in (y_0, x_0]$, we need the condition $y_0 \geqslant \Gamma_y(x_0, u_0^*(x_0)) = \underline{y}_0$. In sum, the incentive conditions are satisfied if $y_0 \geqslant \underline{y}_0$.

Step 3: Verify that the current mover prefers moving to y_k to all actions below.

Step 1 has shown that for any $s \geqslant 0$, setting new state within (y_{k+s+1}, y_{k+s}) is dominated by setting y_{k+s}, so we only need to show that the mover prefers y_{k+s} to y_{k+s+1} for any $s \geqslant 0$. To see this, note

$$\begin{aligned} & U(y_{k+s};x) \geqslant U(y_{k+s+1};x) \\ \Leftrightarrow\quad & \delta r(u_{k+s} - u_{k+s+1}) - (1-\delta)(y_{k+s} - y_{k+s+1})F(x) \\ & +[Lr+\delta(x-y_{k+s})]F(y_{k+s}) - [Lr+\delta(x-y_{k+s+1})]F(y_{k+s+1}) \geqslant 0 \end{aligned}$$

which holds with equality when $x = y_{k+s}$ by (24) and (25). Taking the derivative of the left hand side w.r.t. x gives

$$\begin{array}{rl} & \delta[F(y_{k+s}) - F(y_{k+s+1})] - (1-\delta)(y_{k+s} - y_{k+s+1})f(x) \\ \geqslant & [\delta \underline{f} - (1-\delta)f(x)](y_{k+s} - y_{k+s+1}) \\ > & 0 \end{array}$$

when δ is close to 1 (because of the Lipschitz continuity of f). Hence for any $s \geqslant 0$, $U(y_{k+s}; x) \geqslant U(y_{k+s+1}; x)$ whenever $x \geqslant y_{k+s}$, and by telescoping $U(y_k; x) \geqslant U(y_{k+s}; x)$ for any $s \geqslant 1$.

The three steps above confirm that choosing the new state at y_k is indeed globally optimal when $x \in [y_k, y_{k-1}]$, $\forall k \geqslant 0$, and hence the proposed strategy profile consists a MPE. Moreover, given any $y_0 \in [\underline{y}_0, x_0]$, the sequence $\Gamma^{(n)}(y_0, u_0^*(y_0))$ uniquely pins down a MPE. Hence there is a continuum of MPEs. ■

A.7 Proof of Proposition 11

Proof. We prove the theorem by construction. First we prove that

$$\lim_{\Delta \to 0} \left(\frac{y_{\lfloor \frac{t}{\Delta} \rfloor + 1} - y_{\lfloor \frac{t}{\Delta} \rfloor}}{\Delta} + \frac{\int_{x^*}^{y_{\lfloor \frac{t}{\Delta} \rfloor}} \big([F(s) - F(m)] - 2Lrf(s) \big) ds}{Lf(y_{\lfloor \frac{t}{\Delta} \rfloor})} \right) = 0$$

To see this, note from (25) that

$$\begin{array}{rcl} y_{k+1} - y_k & \geqslant & -\dfrac{(1-\delta^2) r u_k}{\delta[Lr - (1-\delta)x_0][f(y_k) - \kappa(y_k - y_{k+1})]} \\ & \geqslant & -\dfrac{(1-\delta^2) r u_k}{\delta[Lr - (1-\delta)x_0][f(y_k) - (1-\delta)B]}, \end{array} \tag{35}$$

$$\begin{array}{rcl} y_{k+1} - y_k & \leqslant & -\dfrac{(1-\delta^2) r u_k}{\delta Lr[f(y_k) + \kappa(y_k - y_{k+1})]} \\ & \leqslant & -\dfrac{(1-\delta^2) r u_k}{\delta Lr[f(y_k) + (1-\delta)B]} \end{array} \tag{36}$$

where $B = \frac{2r\bar{u}\kappa}{Lr\underline{f}}$.

Meanwhile, note from (24) that $u_{k+1} - u_k = \frac{1-\delta^2}{\delta^2}u_k + \frac{1}{\delta r}(y_{k+1} - y_k)F(y_{k+1})$ so

$$\begin{aligned}
& \frac{F(y_{k+1}) - Lr[f(y_k) + (1-\delta)B]}{\delta r} \\
\leqslant \quad & \frac{u_{k+1} - u_k}{y_{k+1} - y_k} = \frac{1-\delta^2}{\delta^2(y_{k+1} - y_k)}u_k + \frac{1}{\delta r}F(y_{k+1}) \\
\leqslant \quad & \frac{F(y_{k+1}) - [Lr - (1-\delta)x_0][f(y_k) - (1-\delta)B]}{\delta r}
\end{aligned}$$

Hence,

$$\begin{aligned}
& \sum_{s=0}^{\infty} \frac{F(y_{k+s+1}) - Lr[f(y_{k+i}) + (1-\delta)B]}{\delta r}(y_{k+s} - y_{k+s+1}) \\
\leqslant \quad & u_k = u_\infty - \sum_{s=0}^{\infty}(u_{k+s+1} - u_{k+s}) \\
\leqslant \quad & \sum_{s=0}^{\infty} \frac{F(y_{k+s+1}) - [Lr - (1-\delta)x_0][f(y_{k+s}) - (1-\delta)B]}{\delta r}(y_{k+s} - y_{k+s+1}).
\end{aligned}$$

Plugging the above into (35) and (36), we have

$$\begin{aligned}
\frac{y_{\lfloor \frac{t}{\Delta} \rfloor} - y_{\lfloor \frac{t}{\Delta} \rfloor + 1}}{\tilde{\Delta}} &\leqslant \frac{r(1+\delta)}{\delta^2 [Lr - (1-\delta)x_0][f(y_{\lfloor \frac{t}{\Delta} \rfloor}) - (1-\delta)B]} \times \\
& \sum_{s=0}^{\infty} \Big[F(y_{\lfloor \frac{t}{\Delta} \rfloor + s+1}) - [Lr - (1-\delta)x_0][f(y_{\lfloor \frac{t}{\Delta} \rfloor + s}) - (1-\delta)B]\Big](y_{\lfloor \frac{t}{\Delta} \rfloor + s} - y_{\lfloor \frac{t}{\Delta} \rfloor + s+1}), \quad (37) \\
\frac{y_{\lfloor \frac{t}{\Delta} \rfloor} - y_{\lfloor \frac{t}{\Delta} \rfloor + 1}}{\tilde{\Delta}} &\leqslant \frac{r(1+\delta)}{\delta^2 Lr[f(y_{\lfloor \frac{t}{\Delta} \rfloor}) + (1-\delta)B]} \times \\
& \sum_{s=0}^{\infty} \Big[F(y_{\lfloor \frac{t}{\Delta} \rfloor + s+1}) - Lr[f(y_{\lfloor \frac{t}{\Delta} \rfloor + s}) + (1-\delta)B]\Big](y_{\lfloor \frac{t}{\Delta} \rfloor + s} - y_{\lfloor \frac{t}{\Delta} \rfloor + s+1}). \quad (38)
\end{aligned}$$

which can be further simplified to

$$\begin{aligned}
& \frac{2}{Lf(y_{\lfloor \frac{t}{\Delta} \rfloor})} \sum_{s=0}^{\infty} [F(y_{\lfloor \frac{t}{\Delta} \rfloor + s}) - Lrf(y_{\lfloor \frac{t}{\Delta} \rfloor + s})](y_{\lfloor \frac{t}{\Delta} \rfloor + s} - y_{\lfloor \frac{t}{\Delta} \rfloor + s+1}) + C_H(1-\delta) - g(y_{\lfloor \frac{t}{\Delta} \rfloor}) \\
\geqslant \quad & \frac{y_{\lfloor \frac{t}{\Delta} \rfloor} - y_{\lfloor \frac{t}{\Delta} \rfloor + 1}}{\tilde{\Delta}} - g(y_{\lfloor \frac{t}{\Delta} \rfloor}) \\
\geqslant \quad & \frac{2}{Lf(y_{\lfloor \frac{t}{\Delta} \rfloor})} \sum_{s=0}^{\infty} [F(y_{\lfloor \frac{t}{\Delta} \rfloor + s}) - Lrf(y_{\lfloor \frac{t}{\Delta} \rfloor + s})](y_{\lfloor \frac{t}{\Delta} \rfloor + s} - y_{\lfloor \frac{t}{\Delta} \rfloor + s+1}) - C_L(1-\delta) - g(y_{\lfloor \frac{t}{\Delta} \rfloor})
\end{aligned}$$

for some constants C_L and C_H, when δ is close to 1. Also, $g(y) \equiv 2\int_{x_1^*}^{y} \frac{F(s)-Lrf(s)}{Lf(y)} ds$. Note that since $F(x) - Lrf(x)$ is integrable and $y_k - y_{k+1}$ uniformly converges to 0, the left and right hand side both converge to 0 by Dominated Convergence Theorem, and the convergence is uniform in t. By Sandwich Theorem, we know that $\lim_{\Delta\to 0}\left(\frac{y_{\lfloor \frac{t}{\Delta}\rfloor} - y_{\lfloor \frac{t}{\Delta}\rfloor+1}}{\tilde{\Delta}} - g(y_{\lfloor \frac{t}{\Delta}\rfloor})\right) = 0$, uniformly for all $t \geqslant 0$. For a fixed t, $\lfloor \frac{t}{\Delta}\rfloor$ is finite, so $g(y_{\lfloor \frac{t}{\Delta}\rfloor}) > 0$. Also, $\lim_{\Delta\to 0}(\tilde{\Delta} - \Delta) = \lim_{\Delta\to 0} \frac{1-e^{-r\Delta}}{r} - \Delta$=0. The convergence therefore can be rewritten as $\lim_{\Delta\to 0} \frac{y_{\lfloor \frac{t}{\Delta}\rfloor} - y_{\lfloor \frac{t}{\Delta}\rfloor+1}}{g\left(y_{\lfloor \frac{t}{\Delta}\rfloor}\right)\Delta} = 1$. The convergence is now uniform for all times $t' \leqslant t$, meaning that by Sandwich Theorem again,

$$
\begin{aligned}
&\lim_{\Delta\to 0}\left(\left\lfloor \frac{t}{\Delta}\right\rfloor\right)^{-1} \sum_{k=0}^{\lfloor \frac{t}{\Delta}\rfloor-1} \frac{y_k - y_{k+1}}{g(y_k)\Delta} = 1 \\
\Rightarrow \quad &\lim_{\Delta\to 0} \sum_{k=0}^{\lfloor \frac{t}{\Delta}\rfloor-1} \frac{y_k - y_{k+1}}{g(y_k)} = t
\end{aligned}
$$

On the other hand, by Dominated Convergence Theorem we have

$$
0 = \lim_{\Delta\to 0}\left(\sum_{k=0}^{\lfloor \frac{t}{\Delta}\rfloor-1} \frac{y_k - y_{k+1}}{g(y_k)} - \int_{y_{\lfloor \frac{t}{\Delta}\rfloor}}^{x_0} \frac{1}{g(s)} ds\right) = t - \lim_{\Delta\to 0} \int_{y_{\lfloor \frac{t}{\Delta}\rfloor}}^{x_0} \frac{1}{g(s)} ds.
$$

Note that the state at real time t is $x(t;\Delta) = y_{\lfloor \frac{t}{\Delta}\rfloor}$, so that

$$
\lim_{\Delta\to 0} \int_{x(t;\Delta)}^{x_0} \frac{1}{g(s)} ds = \int_{\lim_{\Delta\to 0} x(t;\Delta)}^{x_0} \frac{1}{g(s)} ds = t.
$$

Comparing the above with the time path in continuous time: $\int_{x(t)}^{x_0} \frac{1}{2\nu(s)} ds = t$, and noting that $g(s) = 2\nu(s)$, we immediately have $\lim_{\Delta\to 0} x(t;\Delta) = x(t)$. ■

A.8 Proof of Lemma 12

Proof. For the first part, note that if the breakdown occurs during $t \in (t_a, t_b)$, then it is Player j who incurs the lumpy cost, not Player i. Hence, in the continuation game where

$t = t_a$, $x_i(t_a) = x_j(t_a)$, the present discounted payoff of Player i following path $x_i(\cdot)$ is

$$\begin{aligned} W_i^h(x_i(\cdot), t_a) &= \frac{e^{rt_a}}{F(x(t_a))} \int_0^{x_i(t_b)} \left[\int_{t_a}^{t_b} e^{-rt}(p - x_i(t))dt + e^{-rt_b} W_i^h(x_i(\cdot), t_b) \right] dF(c), \\ &+ \frac{e^{rt_a}}{F(x(t_a))} \int_{x_i(t_b)}^{x_i(t_a)} \int_{t_a}^{t_b} \left[\int_{t_a}^{t} e^{-rs}(p - x_i(s))ds + e^{-rt} U(x_i(t), t) \right] dP_j(t|c) dF(c) \end{aligned}$$

where $U(x_i(t), t)$ is the value function of a (surviving) player whose current state is $x_i(t)$ and her opponent just triggered the threshold at time t. Since $x_i(t) > \tilde{x}_i(t)$ for $t \in (t_a, t_b)$, we have $U(x_i(t), t) < U(\tilde{x}_i(t), t)$ for $t \in (t_a, t_b)$. Note also that $W_i^h(x_i(\cdot), t_b) = W_i^h(\tilde{x}_i(\cdot), t_b)$. Therefore $W_i^h(x_i(\cdot), t_a) < W_i^h(\tilde{x}_i'(\cdot), t_a)$. Since it is with positive probability that time t_a is reached before triggering the breakdown, this inequality carries over to the payoff evaluated at time 0.

For the second part, note that if the breakdown occurs during $t \in (t_a, t_b)$, then it is Player i who incurs the lumpy cost, not Player j. Hence, in the continuation game where $t = t_a$, $x_i(t_a) = x_j(t_a)$, the present discounted payoff of Player i following path $x_i(\cdot)$ is

$$\begin{aligned} & W_i^l(x_i(\cdot), t_a) \\ = \; & -\frac{e^{rt_a}}{F(x(t_a))} \int_{t_a}^{t_b} \left[\int_{t_a}^{t} e^{-rs}(p - x_i(s))ds + e^{-rt} \left(\frac{p - x(t)}{r} - L \right) \right] dF(x(t)) \\ & + \frac{e^{rt_a}}{F(x(t_a))} \left[\int_{t_a}^{t_b} e^{-rt}(p - x(t))dt + e^{-rt_b} W_i^l(x_i(\cdot), t_b) \right] \\ = \; & \frac{e^{rt_a}}{F(x(t_a))} \int_{t_a}^{t_b} e^{-rt} \left(LrF(t) - \int_0^t F(s)ds \right) dt + \frac{e^{rt_a}}{F(x(t_a))} F(x(t_b)) e^{-rt_b} W_i^l(x_i(\cdot), t_b) \\ & + \frac{e^{rt_a}}{rF(x(t_a))} \left[e^{-rt_a} \int_0^{x(t_a)} F(s)ds - e^{-rt_b} \int_0^{x(t_b)} F(s)ds \right] \\ & + \frac{e^{rt_a}}{rF(x(t_a))} \left[e^{-rt_a} F(x(t_a))(p - Lr - x(t_a)) - e^{-rt_b} F(x(t_b))(p - Lr - x(t_b)) \right], \end{aligned}$$

Since $x_i(t_a) = \tilde{x}_i(t_a)$, $x_i(t_b) = \tilde{x}_i(t_b)$, the only term that matters is the first one (with the integral). Pointwise differentiation inside the integral gives

$$\frac{\partial}{\partial x(t)} e^{-rt} \left(LrF(t) - \int_0^t F(s)ds \right) = e^{-rt}(Lrf(x(t)) - F(x(t))) < 0, \tag{39}$$

so that $W_i^l(x_i(\cdot), t_a) < W_i^l(\tilde{x}_i'(\cdot), t_a)$. Since it is with positive probability that time t_a is reached before triggering the breakdown, this inequality carries over to the payoff evaluated at time 0. ■

A.9 Proof of Proposition 13

Proof. Suppose there were a pure strategy Markov equilibrium that conditional on no breakdown induces paths of effort $x_1(\cdot)$ and $x_2(\cdot)$, respectively. Define $t_a \equiv \inf\{t \in \mathbb{R}_+ : x_1(t) \neq x_2(t)\} \in \mathbb{R}_+ \cup \{\infty\}$.

If $t_a = \infty$, then $x_1(t) = x_2(t)$ for all $t \geqslant 0$. There are two subcases: $x_1(t) = x_2(t) = x_0$ for all $t \geqslant 0$, or not. In the former subcase, either player has incentive to deviate to the path for single player (see Section 3) because $\frac{F(x_0)}{f(x_0)} > Lr$ by assumption. In the latter subcase, either player, say Player i, could deviate to a feasible path $\tilde{x}_i(t) = x_i(\max\{t - \varepsilon, 0\})$ for some small $\varepsilon > 0$, paying slightly more flow cost than before while avoiding breakdown completely.

If $t_a < \infty$, then define $t_b \equiv \sup\{t > t_a : x_1(t) \neq x_2(t) \text{ on } (t_a, t_b)\} \in (t_a, \infty) \cup \{\infty\}$. By continuity and differentiability of $x_1(\cdot)$ and $x_2(\cdot)$, we know that $x_1(t_a) = x_2(t_a)$, $x_1(t_b) = x_2(t_b)$ (if $t_b < \infty$), and either $x_1(t) > x_2(t)$ for all $t \in (t_a, t_b)$ or $x_1(t) < x_2(t)$ for all $t \in (t_a, t_b)$. Without loss of generality suppose $x_1(t) > x_2(t)$ for $t \in (t_a, t_b)$. If $t_b < \infty$, then apply Lemma 12 to see that

$$\tilde{t}_1(t) = \begin{cases} \alpha t_1(t) + (1-\alpha) t_2(t) & \text{if } t \in (t_a, t_b) \\ t_1(t) & \text{otherwise} \end{cases}$$

is a profitable deviation for Player 1, where $\alpha \in (0, 1)$. ■