# Miscoordination and Delay in Strategic Experimentation

Alessandro Bonatti[*], Johannes Hörner[†]

April 2014

**Abstract**

We analyze strategic experimentation where information arrives through fully revealing, publicly observable "breakdowns." We contrast the case of observable and unobservable actions. In the unobservable case, there exists a unique symmetric equilibrium that (for some parameters) involves randomization over stopping times. The timing of experimentation and overall welfare are worse than under observability. Taking monitoring as a design variable, this may explain information-sharing practices and regulations in several industries. Taking lack of observability as a constraint, we explore potential remedies, such as subsidies and risk-sharing.

[*]MIT Sloan School of Management, 100 Main Street, Cambridge MA 02412, USA `bonatti@mit.edu`.
[†]Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, `johannes.horner@yale.edu`.

# 1 Introduction

Recent empirical literature has documented a wide dispersion in the practices adopted by different firms within an industry and the resulting differences in productivity. While dispersion in productivity cannot be solely traced to heterogeneity in practices, an important question is to evaluate the role of new technologies on firm performance.

Nearly every study attempting to document the dispersion in performance across several firms has used a single agent and either (a) path dependence, (b) multiplicity stemming from repeated interaction, or (c) heterogeneity in parameters with tipping points. Very few, if any, look at the role of market interaction for the diffusion of technologies.

We examine the role of informational externalities on the adoption of new technology. We show that heterogeneity in technology adoption emerges as the unique (symmetric) equilibrium outcome of strategic experimentation. We abstract from product-market competition and consider firms operating in local monopolies. Thus, all firms benefit equally from information, but they do not share equally into costly breakdowns. This allows us to isolate the free-riding and encouragement effects associated with strategic experimentation in an industry. We quantify the value of information, which refers both to the unknown technological state and to the ability to observe other agents' actions.

More concretely, we seek to characterize: (i) the distribution of adopters and the patterns of diffusion; (ii) the welfare consequences of opaque vs. transparent settings; (iii) the benefits of coordination (correlated equilibrium) as an industry benchmark; and (iv) the potential role of subsidies and government intervention (e.g., insurance against harvest loss for technology adoption in agriculture; or group risk-sharing).

Examples include the adoption of new risky technologies. Within our framework, a main theme will be "too little experimentation" (use of the risky arm). This is not to be confused with cross-market externalities. For instance, in the energy sector, ignoring the environmental impact of a breakdown may well yield excessive risk-taking by the industry as a whole.

We are not limited to new risky technological advances. For instance, one may view the use of state-of-the-art managerial practices as a very specific new technology. In some cases, modern manufacturing standards can be applied and yield unambiguous benefits. In other instances (e.g., developing countries) the match of new practices with local conditions is much less clear.

Add discussions closer to the role of monitoring: rules pro-transparency in the fracking procedure for drilling oil and gas; "best practices" and "industry associations" as means to share information, e.g. among hospitals.

**Mixing:** The necessity of mixed strategies should not be confused with the necessity of using actions that are not extremal. To restore existence in games of strategic experimentation without having to confront the measure-theoretic difficulties raised by the modeling of independent randomizations in continuous time, various authors (see Bolton and Harris 1999; Keller, Rady and Cripps, 2005; Keller and Rady, 2014, in particular) have redefined the space of actions available to a player at a given instant to be a convex set (usually replacing the lotteries over $\{0,1\}$ by the interval $[0,1]$). In all these papers, this suffices to restore equilibrium existence (relative to the decision-theoretic versions of bandit problems, which admit optimal solutions within the class of extremal policies, see Yushkevich, 1988, or Presman and Sonin, 1990).

Similarly, while standard games or pre-emption or wars of attrition are usually described in terms of games of timing, in which the randomization is over the stopping time (indeed, unlike games of experimentation, these are games in which all actions but one terminate the game), the issue of existence would be resolved in those papers by allowing action sets that are sectionally convex (that is, that are convex at every instant of time). Considering, say, the war of attrition of Milgrom and Weber (1985), or more recent versions of timing games allowing for additional learning (Murto and Välimäki, 2011; Rosenberg, Salomon and Vieille, 2013) it is a matter of interpretation whether the hazard rate used by the players in their equilibrium strategy is an *ex ante* randomization over stopping times or (a map from time into) action(s) that takes arbitrary values in $[0,1]$. While Kuhn's theorem obviously carries over to our environment as well (see Weizsäcker, 1974, or Shmaya and Solan, 2014), there is an important difference between these papers and ours: in all these papers, a player is made indifferent over *all* his strategies (over the relevant time interval). In our model, a player is indifferent over stopping times, as long as these times belong to a given interval, but he strictly prefers any of these stopping times to a strategy that would use an interior action over a set of times of positive measure. In all these papers, by redefining action spaces so as to make them sectionally convex, existence of equilibrium (in pure strategies) is restored; not in ours.

Of course, it is well known in optimization that sectionally convexity is insufficient to guarantee the kind of convexity in the strategy space that is required for existence of solutions of optimal control problems. *A fortiori*, the problem arises in games, and there are well-known examples of zero-sum games with sectionally convex action spaces for which the optimal strategies cannot be found within the class of pure strategies (see Karlin, 1959 and references therein). However, we are not aware of another paper with a clear economic interpretation that would display the same property. Our paper shows that such phenomena are both relevant for economic applications and amenable to mathematical analysis. (See

Board and Meyer-ter-Vehn, 2014, for another model in which mixing might, or might not be required, as existence of pure-strategy equilibria is not established. See also Thomas, 2014.)

**Uniqueness:** Another feature that distinguishes this game from related ones is the uniqueness of equilibrium. Not only is the symmetric equilibrium unique (and for some parameters, as mentioned, mixed), there is no asymmetric pure equilibrium, or for that matter (at least for many parameters in the two-player case, and we suspect for all) no asymmetric mixed-strategy equilibrium either. Our game does not feature complementarities, and indeed, because of the non-convexity of the best-reply problem to a pure strategy by the opponent which causes the equilibrium to be mixed, best-reply functions (to pure strategies) are not monotone. They are first increasing, then decreasing, with a jump downward in between. This implies that standard methods to prove uniqueness fail.[1] Our proof carries no philosophical charm, and is based on particular features of the payoff function.

Uniqueness contrasts with the multiplicity that is prevalent in games with strategic experimentation, not only when actions are observable (Keller, Rady and Cripps, 2005; Keller and Rady, 2014), but also when they are not (Bonatti and Hörner, 2011). Because of the pervasive free-riding incentives, asymmetric equilibria typically exist, where players alternate (finitely or infinitely often) between experimenting and taking advantage of the opponent's experimentation –leading to the existence of additional asymmetric equilibria. By contrast, in our game, free-riding finds its expression in how early a player is willing to start experimenting; the earlier the opponent does, the later one finds it optimal to do so. But the *ordering* of actions is unambiguous: for a given total amount of experimentation, it is always best to use a stopping strategy, using the safe arm if and only if a threshold time has not yet been reached.[2] It is impossible to give a player's incentives to use a strategy that would involve using the risky arm before the safe arm, precluding any type of alternation in the experimentation that players conduct.

**Observability:** Finally, a feature that distinguishes our model from the literature is that monitoring (that is, observing the other players' actions) is good. That is, the equilibrium

[1]See Vives (1999) for an excellent discussion of these. The jump down in the best-reply curves makes existing arguments based on supermodular games ineffective; the best-reply function is not a contraction either (the equilibrium would be pure otherwise), and the fact that the equilibrium is mixed implies that the Gale-Nikaido theorem or the Poincaré-Hopf theorem cannot work either, or rather, one should work with the mixed-strategy space directly, and possibly use an infinite-dimensional extension of those.

[2]This is also the key reason for why equilibrium must be in mixed strategies, and not in pure strategies with non-extremal actions: for a given amount of experimentation, players have strict incentives to backload it, ruling out the use of arms with non-extremal intensity over any interval of time.

payoff is higher in the symmetric Markov equilibrium with observability than in the game without such observability. This is true despite the fact that attention is restricted to symmetric Markov, ruling out the type of threats and punishments that observability could enable and that would clearly make cooperation easier to sustain. In contrast, monitoring is harmful in, say, Bonatti and Hörner (2011) and related games with incomplete information.[3] The reason behind the difference is explained in greater detail in Section 6, but the basic intuition is easy to grasp: when actions are observable, a player's incentive to deviate are related not only to the direct cost or benefit from this deviation, but also the indirect cost or benefit in terms of the change of actions by the other players. By deviating to the risky arm, a player accelerates the common learning, which, in the absence of news, leads to more optimism and more experimentation by others; this is good, because players do not experiment enough. In contrast, with good news, experimentation by a player leads to more pessimism in the absence of news, and hence depresses experimentation provision.

**Policy:** Finally, we show that a well-chosen "insurance" scheme across players, whereby a player who suffers a breakdown gets a partial compensation from all other players can achieve first-best. Remarkably, the appropriate amount of compensation is the same whether actions are monitored or not. That is, the precise monitoring opportunities of players turn out to be irrelevant. We explain this result and the intuition behind it in Section 6.

# 2 The Game

## 2.1 Set-Up

Time is continuous and the horizon is infinite. Players $i = 1, \ldots, I$ ($I \geq 2$) choose an effort level (or action) $u^i \in [0, 1]$ at all times.

There is a binary state of the world $\omega \in \{B, G\}$. Players assign a common prior probability $p^0 \in (0, 1)$ to the event $\{\omega = B\}$. Conditional on $\omega$, player $i$'s action controls the instantaneous intensity of a conditionally independent Poisson process $\{N_t^i : t \geq 0\}$, where $N_t^i$ denotes the number of lump-sums observed up to time $t$. That is, the effort paths $u^i = (u_t^i)_{t=0}^{\infty}$, alongside $\omega$, define the instantaneous intensity of an inhomogeneous Poisson process with intensity $\lambda(t) := \lambda \mathbf{1}_{\{\omega=B\}}(\bar{u}/I - u_t^i)$, where $\bar{u} \geq I$, $\lambda > 0$ and $\mathbf{1}_A$ is the indicator function of an event $A$. Note that this intensity is zero if $\omega = G$, independently of the effort levels chosen. When $u^i = 1$, player $i$ uses exclusively the *safe arm*, as this maximally reduces

[3] Holmström, 1999, being perhaps the most famous such example, though arguably the mechanism through which lack of observability operates there is very different.

the probability of (costly) lump-sums. This does not imply that lump-sums cannot occur in that case, unless $\bar{u} = I$, a special case of particular interest. The case $\bar{u} > I$ will be referred to as the game with *background learning*, as beliefs evolve even when all players use the safe arm. When player $i$ sets $u^i$ to 0, we say that he uses exclusively the *risky arm*.

Each lump-sum entails a cost $h > 0$. That is, given a (measurable) function $u = (u^i_t)$, and the realization of the process $\{N^i_t : t \geq 0\}$, the realized cost of player $i$ is given by

$$\int_0^\infty re^{-rt}(h\mathrm{d}N^i_t + su^i_t\mathrm{d}t),$$

where $r, s > 0$. The case in which players' parameters are not identical is discussed in Section 5. Note that this is a game of informational externalities only, as player $j \neq i$'s actions do not enter player $i$'s cost.

We assume that $g := \lambda h > s$: therefore, conditional on $\{\omega = B\}$, to minimize the cost it is optimal to allocate the resource exclusively to the safe arm, that is, to set $u^i_t = 1\ \forall t$. Conditional on $\{\omega = G\}$, the risky arm is optimal. Throughout, we assume that player $i$ observes the collection of processes $\{N^i_t : i \in I\}$ and can condition his action on it. Hence, player $i$'s problem reduces to a choice of action up to the first arrival, as it is strictly dominant to use the risky arm afterwards. Let $\tau \in \mathbf{R}_+ \cup \{+\infty\}$ be the time of first arrival. (Note that $\tau = +\infty$ if $\omega = G$.) Hence, we are interested in the interaction on the horizon $[0, \tau]$ only. A terminal history $h^\tau$ specifies the stopped effort paths $\{(u^i_t)^\tau_{t=0} : i = 1, \ldots, I\}$ up to time $\tau$. We can rewrite the cost whose expectation is to be minimized as

$$\mathcal{C}^i(u^i) = \int_0^\tau \left(re^{-rt}su^i_t\mathrm{d}t + e^{-rt}rh\mathrm{d}N^i_t\right) + e^{-r\tau}\bar{K}, \tag{1}$$

where

$$\bar{K} := \left(\frac{\bar{u}}{I} - 1\right)g + s$$

is a terminal cost equal to the expected value conditional on $\{\omega = B\}$ and $u^i_t = 1\ \forall t$.

Some of the parameters matter in combination only. In particular, up to a normalization, $g$ and $s$ only enter through the cost-benefit ratio $\gamma := (g - s)/s$, and, via a standard change, the discount rate $r$ and intensity parameter $\lambda$ only appear via the ratio $\mu := r/\lambda$.

We assume that player $i$ observes nothing but the values of these processes $\{N^i_t : i \in I\}$. In particular, player $i$ does not observe the past values of $u^j_t$, $j \neq i$.

## 2.2 Strategies

A deterministic policy (or strategy) for player $i$ is a measurable function $\pi^i : \mathbf{R}_+ \to [0, 1]$ that specifies player $i$'s effort $u^i$ at time $t$ conditional on the event $\{t < \tau\}$. Let $\Pi^i$ denote the set of all deterministic policies.

It is not enough to consider deterministic policies, as it turns out. Mixed policies must be considered, calling for additional notation. Let $\mathcal{B}_{[0,1]}$ (resp. $\mathcal{B}$) denote the $\sigma$-algebra of Borel sets of $[0,1]$ (resp. $\mathbf{R}_+$) and $\lambda$ denote the Lebesgue measure on $[0,1]$. We endow the set of measurable functions from $(\mathbf{R}_+, \mathcal{B})$ to $([0,1], \mathcal{B}_{[0,1]})$ with the $\sigma$-algebra generated by sets of the form $\{f : f(s) \in A\}$ with $s \in \mathbf{R}_+$ and $A \in \mathcal{B}_{[0,1]}$. We adopt the following definition of mixed policies due to Aumann (1964). A mixed policy is a measurable map $\phi^i : [0,1] \to \Pi^i$ such that for all $\beta^i \in [0,1]$, $\phi^i(\beta^i) \in \Pi^i$. (That such a definition is equivalent to the use of "behavioral decision rules" follows from Weizsäcker, 1974. See also Shmaya and Solan, 2014 for this equivalence, and Touzi and Vieille, 2002 on randomized strategies in timing games.) Let $\Phi^i$ denote the set of (mixed) policies of player $i$.

Given $\phi^{-i} \in \Phi^{-i} := \times_{j \neq i} \Phi^j$, player $i$ minimizes

$$\mathcal{C}^{\phi^i} := \mathbf{E}_{p^0}^{\phi^i}\left[\mathcal{C}^i(u^i)\right],$$

over $\phi^i \in \Phi^i$.

Given $t \geq 0$, we write $\pi_t^i$ for the strategy $\pi_t^i(s) = 1$ for $s < t$, $\pi_t^i(s) = 0$ for $s \geq t$. The set of such stopping policies is denoted $\Pi_S^i$. Of particular interest are *stopping time* strategies. These are policies such that, for some non-decreasing function $t^i : [0,1] \to \mathbf{R}_+$, $\phi^i(\beta^i) = \pi_{t^i(\beta^i)}^i$ ($\lambda$-a.s.). In words, these are strategies in which player $i$ randomizes over the time he stops exerting high effort. Let $\Phi_S^i$ denote the set of stopping time strategies of player $i$ (whether pure or randomized). It will be more convenient to represent such strategies by the distribution function $F^i : \mathbf{R}_+ \to [0,1]$ defined as $F^i(t) = \sup\{\beta^i \mid t^i(\beta^i) \leq t\}$ (that is, $t^i$ is the quantile function of $F^i$).

Note that, given that players do not observe each others' actions, there is no loss in considering Nash equilibria. Hence, an equilibrium is a vector $\phi^* \in \Phi := \times_i \Phi^i$ such that, for all $i$ and for all $\beta^i \in [0,1]$, $\phi^{*i}(\beta^i)$ minimizes $\mathcal{C}^{\phi^i}$ over $\phi^i \in \Phi^i$, given $\phi^{*-i}$. Of particular interest are symmetric equilibria, that is, equilibria in which $\phi^j = \phi^i$ for all $i, j$.

## 2.3 Cooperative Solution

Let us first solve the cooperative problem. Assume that players perfect observe each others' action and choose them so as to maximize the sum of of their costs. As we shall see, because a profile of pure strategies is optimal, implementing the solution does not require perfect monitoring, so that it all applies to the case in which players maximize joint cost without such observation.

As is standard, we reduce the stochastic problem to a deterministic one by introducing the random process of posterior probabilities $p_t := \mathbf{P}_{p^0}^{\phi}[\omega = B \mid (u_s^i)_{s=0}^t, i = 1, \ldots, I]$, $p_0 = p^0$, conditional on the event $\varnothing_t := \{\tau > t\}$. Also, it is convenient to work with the equivalent process of log-likelihood (or "odds") ratios $(\ell_t)_{t\geq 0}$, with

$$\ell_t := \ln p_t/(1-p_t),$$

and we set $\ell^0 := p^0/(1-p^0)$. Bayes' rules gives that, up to the first breakdown, the process $\ell$ follows the differential equation

$$\dot{\ell}_t = \sum_i u_t^i - \bar{u}, \quad \ell_0 = \ell^0.$$

Define

$$\ell^{FB} := \ln \frac{\mu + \bar{u}}{\mu + \bar{u} - I} \frac{1}{\gamma},$$

and define $(\ell_t^{FB})_{t\geq 0}$ by $\ell_t^{FB} = \ell^0 + (I - \bar{u})t$ as long as $t < t^*$, and $\ell_t^{FB} = \ell^0 + It^* - \bar{u}t$ for $t \geq t^*$, where $t^*$ solves $\ell^0 + (I - \bar{u})t^* = \ell^{FB}$. The time $t^*$ is the delay that it takes for the odds ratio to reach $\ell^{FB}$, when all players use the safe arm from time 0 to $t^*$.

Given a pair $(\ell, u)$ such that $\ell$ is the belief path generated by $u := \sum_i u^i$ given $\ell^0$, along the history with no breakdown, the effort path $(u_t)_t$ is *measurable with respect to the belief path* $(\ell_t)_t$ if $\ell_t = \ell_{t'} \Rightarrow u_t = u_{t'}$ for all $t, t'$. Effort paths are always measurable with respect to the belief path whenever $\bar{u} > I$, as independently of the actions chosen, no two distinct times get mapped into the same belief. We write $u(\ell)$ for the value of $u$ at the belief $\ell \leq \ell^0$, which is well-defined. The cooperative solution given in the next lemma is measurable with respect to its belief path.

Unsurprisingly, the optimal policy consists in having all players apply the safe arm until belief $\ell^{FB}$ is reached, and then switch to the risky arm. This is stated below (see also Keller and Rady 2014, Proposition 1). The next lemma also that welfare decreases in the intensity with which the risky arm is used, as long as the ranking of intensities holds pointwise in the beliefs.

**Lemma 1** *The cooperative solution $u^{FB}$ is given by $u_t^{FB} = I$ for all $t$ such that $\ell_t \geq \ell^{FB}$, and $u_t^{FB} = 0$ otherwise. Furthermore, let $\ell', \ell'' : \mathbf{R}_+ \to \mathbf{R}$ be two feasible paths such that the corresponding effort path $u', u''$ be measurable with respect to their belief path, with for all $\ell \leq \ell^0$, $u^{FB}(\ell) \leq u'(\ell) \leq u''(\ell)$. Then welfare is weakly higher under $\ell'$ than under $\ell''$, and strictly higher whenever $u'(\ell'_t) < u''(\ell'_t)$ for a set of times $t$ of positive measure.*

## 3 Preliminary Results

### 3.1 Reformulation of the Objective

Our first task is to turn each player's problem into a deterministic control problem. This is more difficult than for the social planner, because a player does not only face uncertainty regarding the state of the world, but also regarding the strategy chosen by its rivals, which affects his learning about the state. Note also that players' beliefs may differ, as a player's own past effort choices enters how much he learns. Hence, player $i$'s actual belief is *not* common knowledge among players, unless player $i$ uses a pure strategy.

Throughout, fix a player $i$. Define $p_t^i := \mathbf{P}_{p^0}^{\phi}[\omega = B \mid (u_s^i)_{s=0}^t]$; as the conditioning makes clear it is player $i$'s belief and his only, although we will occasionally drop the subscript. Two properties of his belief process are important for the sequel. First, it gives a convenient measure of the probability of the event $\varnothing_t$ (that no breakdown has occurred yet); namely, by the martingale property of beliefs,

$$\mathbf{P}_{p^0}^{\phi}[\varnothing_t] \cdot p_t^i + (1 - \mathbf{P}_{p^0}^{\phi}[\varnothing_t]) \cdot 1 = p^0,$$

so that

$$\mathbf{P}_{p^0}^{\phi}[\varnothing_t] = \frac{1 - p^0}{1 - p_t^i}.$$

Second, we can derive the law of motion of the (deterministic) process $p_t$, taking into account the possibly stochastic strategy by players $-i$. Given that breakdowns follow an exponential distribution, the probability of no breakdown by time $t$, conditional on $\{\omega = B\}$ and $(u_s^i)_{s=0}^t$, so as evaluated by player $i$, is given by

$$\mathbf{P}_{p^0}^{\phi}[\varnothing_t \mid \omega = B] = \mathbf{E}_{p^0}^{\phi}\left[e^{-\int_0^t \lambda\left(\bar{u} - \sum_j u_s^j\right)\mathrm{d}s} \mid (u_s^i)_{s=0}^t\right].$$

By independence and conditional on $(u_s^i)_{s=0}^t$, this is equal to

$$\mathbf{P}_{p^0}^{\phi}[\varnothing_t \mid \omega = B] = e^{-\int_0^t \lambda(\bar{u} - u_s^i)\mathrm{d}s} \Pi_{j \neq i} \mathbf{E}_{p^0}^{\phi}\left[e^{\int_0^t \lambda u_s^j \mathrm{d}s}\right].$$

Hence, an alternative formula for $p_t^i$ is given by

$$\frac{p_t^i}{1 - p_t^i} = \frac{p^0}{1 - p^0} \frac{\mathbf{P}_{p^0}^{\phi}[\varnothing_t \mid \omega = B]}{\mathbf{P}_{p^0}^{\phi}[\varnothing_t \mid \omega = G]} = \frac{p^0}{1 - p^0} e^{-\int_0^t \lambda(\bar{u} - u_s^i)\mathrm{d}s} \Pi_{j \neq i} \mathbf{E}_{p^0}^{\phi}\left[e^{\int_0^t \lambda u_s^j \mathrm{d}s}\right]. \tag{2}$$

Note that the log-likelihood ratio

$$\ln \frac{\mathbf{P}_{p^0}^{\phi}[\varnothing_t \mid \omega = B]}{\mathbf{P}_{p^0}^{\phi}[\varnothing_t \mid \omega = G]} = \sum_{j \in I} \ln \mathbf{E}_{p^0}^{\phi}\left[e^{\int_0^t \lambda(u_s^j - \bar{u})\mathrm{d}s}\right]$$

is differentiable with respect to $t$, and we define

$$\nu_t^{-i} := \sum_{j \neq i} \frac{1}{\lambda} \frac{\partial}{\partial t} \ln \mathbf{E}_{p^0}^{\phi}[e^{\int_0^t \lambda u_s^j \mathrm{d}s}]. \tag{3}$$

The function $\nu^{-i}$ plays an important role in the analysis. Note that $\nu_t^{-i} \in [0, I-1]$. It can be interpreted as the expected contribution from the other players' experimentation to the hazard rate of player $i$'s belief. Player $j \neq i$'s experimentation affects player $i$'s belief revision at time $t$, and it is not only a matter only of whether player $j$ has stopped playing safe by then, but when he has done so: if player $i$ were to know when this happened, it would affect his belief about the state of the world at time $t$, and hence by how much this belief must be revised if no breakdown occurs in the next instant.

It follows from (2) that $(p_t^i)$ is also differentiable, and that it solves the differential equation

$$\dot{p}_t^i = -\lambda p_t^i (1 - p_t^i)(\bar{u} - u_t^i - \nu_t^{-i}), \quad p_0^i = p^0. \tag{4}$$

Alternatively, defining $\ell_t^i := \ln p_t^i/(1-p_t^i)$ as the log-likelihood ratio of $i$'s belief (here as well, we sometimes omit the subscript $i$), we have

$$\dot{\ell}_t^i = -\lambda(\bar{u} - u_t^i - \nu_t^{-i}), \quad \ell_0^i = \ell^0, \tag{5}$$

recalling that $\ell^0 = p^0/(1-p^0)$. Note that player $i$'s own effort $u^i$ and $\nu^{-i}$ appear separately in player $i$'s belief revision. Hence, while player $i$'s belief is private, the contribution to this revision that can be attributed to the lack of observability is common knowledge.

By the law of iterated expectations, we may now rewrite the problem of minimizing (1) as, equivalently, minimizing

$$\int_{t \geq 0} e^{-rt} \left( r p_t^i g \left( \frac{\bar{u}}{I} - u_t^i \right) + r u_t^i s + \lambda p_t^i (\bar{u} - u_t^i - \nu_t^{-i}) \left( s + \left( \frac{\bar{u}}{I} - 1 \right) g \right) \right) \frac{1-p^0}{1-p_t^i} \mathrm{d}t, \tag{6}$$

over all paths $(u_t^i)_{t \geq 0}$, subject to (4). The interpretation is the following. As explained, $(1-p^0)/(1-p_t^i)$ is the probability of reaching time $t$ without a breakdown. At that time, if player $i$ is investing $u_t^i$ on the safe arm, the instantaneous probability he suffers a breakdown is $(\bar{u}/I - u_t^i)\lambda p_t^i \mathrm{d}t$, with expected cost $rh$. If any of the players has a breakdown (which occurs with probability $\lambda p_t^i(\bar{u} - u_t^i - f_t)\mathrm{d}t$), then everyone switches to the safe arm, yielding a cost of $s + (\bar{u}/I - 1)g$ (in net present value).

Because $u_i \in [0, 1]$, each agent faces an uninsurable risk component, whose expected cost is given by $K := p^0 g(\bar{u}/I - 1)$. We can therefore rewrite the objective as

$$\int_{t \geq 0} e^{-rt} \left( r p_t^i g \left( 1 - u_t^i \right) + r u_t^i s + \lambda p_t^i (\bar{u} - u_t^i - \nu_t^{-i}) s \right) \frac{1-p^0}{1-p_t^i} \mathrm{d}t + K. \tag{7}$$

This formulation is then consistent with two interpretations of our model: (a) if $K > 0$, agents exert effort towards risk reduction, and they cannot fully insure against breakdowns; (b) if $K = 0$, agents choose between a risky and a safe technology and they can fully insure themselves against breakdowns. However, if $\bar{u} > I$ the industry receives "background information" (say, from a competitive fringe), even if all players choose the safe technology. This is captured by the arrival rate $\lambda p_t(\bar{u} - u_t^i - \nu_t^{-i}) > 0$.

This is the deterministic problem that we study. However, we can simplify the problem somewhat: recall that $\gamma := (g - s)/s$ and $\mu := r/\lambda$. With some elementary transformations (see Appendix) we obtain the program $\mathcal{P}$:

$$\inf \int_t e^{-\mu t} \left( \mu \ell_t^i - \gamma(\bar{u} - \nu_t^{-i} - 1 + \mu) e^{\ell_t^i} \right) \mathrm{d}t \tag{8}$$

over all $\pi^i : \mathbf{R}_+ \to [0, 1]$, measurable, subject to

$$\dot{\ell}_t^i = u_t^i + \nu_t^{-i} - \bar{u}, \quad \ell_0^i = \ell^0.$$

Here, the function $\nu^{-i} : \mathbf{R}_+ \to [0, I - 1]$ is treated as an exogenous (measurable) function (in terms of player $-i$'s strategies, it is given by (3), with given the time change, $\lambda = 1$, but we will not need this before Section 4.1). We omit it as an explicit argument of $\mathcal{P}$, but it parametrizes the program nonetheless.

By the theorem of Filippov-Cesari (see Cesari 1983), a solution exists, that is, the infimum is achieved. We will examine the necessary conditions given by Pontryagin's maximum principle. Further, it is easy to see that the program $\mathcal{P}$ is not abnormal (see Seierstad and Sydsæter, 1987, Ch.2.4, note 5).[4]

## 3.2 Stopping Policies

Here we show that any best-reply must be within the class of stopping policies. The following lemma is proved in Appendix.

**Lemma 2** *If $\pi^i \in \Pi^i$ solves $\mathcal{P}$, then $\pi^i \in \Pi_S^i$.*

Informally, Lemma 2 says that if a player starts experimenting, he should do so forever (*i.e.*, until a breakdown occurs), and conversely, if one plays safe, he should have played safe at all earlier times. To gain more intuition, consider the *arbitrage equation* of player $i$, which considers the trade-off between *backloading* and *frontloading* experimentation. This equation

[4]The argument given Seierstad and Sydsæter (1987) must be slightly modified, as it applies to a fixed horizon. The adjustment is straightforward.

does not establish whether agent $i$ wishes to experiment at time $t$, but only his preference over the *timing* of a fixed amount of experimentation. The marginal value of *backloading* experimentation is given by

$$r\left(p_t^i g - s\right) + \lambda p_t^i \left(\bar{u} - u_t^i - \nu_t^{-i}\right)(g-s) - \lambda p_t^i (g-s)\left(1 - u_t^i\right). \tag{9}$$

The first term is the time-preference effect of delaying the expected flow cost $p_t g$ and anticipating the cost $s$. The second term considers the instantaneous probability of a breakdown $\lambda p_t^i \left(\bar{u} - u_t^i - \nu_t^{-i}\right)$: if one occurs at $t$, safe would be played at $t + \mathrm{d}t$ regardless of the agent's earlier action; in that event, playing more safe at $t$ yields marginal savings of $g - s$. Finally, the third term considers the effect of the agent's action on the likelihood of a breakdown: by frontloading safe, the agent reduces (at rate $\lambda p_t^i$) the arrival of a breakdown, in which case he would switch from the current action $u_t^i$ to $u = 1$; because this can only occur in the bad state, this yields a loss $g - s$.

Note that the sum of the last two terms is non-negative. Hence, equation (9) implies that backloading is profitable whenever $p$ is large enough. Conversely, if an agent were sure that the state is good, discounting would suggest frontloading the risky action. Lemma 2 then establishes that over the relevant range of beliefs (*i.e.*, for $\ell^i \geq \ell^*$, see Lemma 3), the marginal value of backloading is positive.

Finally, note that Lemma 2 does not imply that the solution to $\mathcal{P}$ is unique. Rather, it implies that all deterministic solutions are in $\Pi_S^i$. Hence, in the following equilibrium analysis, we can restrict attention to strategy profiles in $\Phi_S$, the set of (possibly randomized) stopping time strategy profiles. Accordingly, we can work with the distributions functions $F^i : \mathbf{R}_+ \to [0,1]$ and the complementary distribution functions $\bar{F}^i = 1 - F^i$. This also allows to provide an alternative, perhaps more expressive formula for $\nu^{-i}$. By definition,[5]

$$\nu_t^{-i} = \sum_{j \neq i} \frac{\partial}{\partial t} \ln \left[ e^t (1 - F_t^j) + \int_0^t e^s \mathrm{d}F_s^j \right],$$

so that, explicitly,

$$\nu_t^{-i} = \sum_{j \neq i} \frac{e^t \bar{F}_t^j}{\bar{F}_0^j + \int_0^t e^s \bar{F}_s^j \mathrm{d}s}. \tag{10}$$

It follows that $\nu_t^{-i}$ is a function that starts at $I - 1$, remaining there as long as $F^j(t) = 0$ for all $j \neq i$, jumps down whenever $F^j$ jumps up for some $j \neq i$, and continuously increases whenever $t \notin \cup_{j \neq i} \operatorname{supp} F^j$, strictly so unless it is equal to 0, which occurs when $F^j(t) = 1$ for all $j \neq i$.[6]

[5]Recall that $\lambda = 1$ given the change of variable leading to $\mathcal{P}$.

[6]Given a distribution $G$, we write $\operatorname{supp} G$ for the set of points of increase of $G$.

### 3.3 An Upper and a Lower Bound to the Use of the Safe Arm

Here we provide a lower and an upper bound to the use of the same arm, in terms of beliefs. We define

$$\ell^* := \ln \frac{\mu + \bar{u}}{\mu + \bar{u} - 1} \frac{1}{\gamma},$$

as well as

$$\ell^{**} := \ln \frac{1 + \mu}{\bar{u} + \mu - I} \frac{1}{\gamma}.$$

The next result states that once a player gets sufficiently optimistic (more precisely, when $\ell_t^i < \ell^*$), he allocates his entire resource to the risky arm. Conversely, some player must assign positive probability to using the safe arm until his belief reaches $\ell^*$.

**Lemma 3** *If $\pi^i$ solves $\mathcal{P}$, then $u_t^i = 0$ for all $t$ such that $\ell_t^i < \ell^*$. Conversely, if $u_t^i = 0$ yet $\ell_t^i > \ell^*$, then $\nu_t^{-i} > 0$.*

**Proof.** Consider the continuation cost corresponding to the objective (8), defined as

$$\mathcal{C}(\ell, t) := \int_{s \geq t} e^{-\mu s} \left( \mu(\ell + \chi_s) + \gamma(\nu_s^{-i} - \bar{u} - \mu + 1) e^{\ell + \chi_s} \right) \mathrm{d}s,$$

where we define $\chi_s := \int_{\tau=t}^{s} (\nu_\tau^{-i} - \bar{u}) \mathrm{d}\tau$ be the value from setting $u^i = 0$ (identically), given $\ell$ and $t$. Note that, integrating by parts,

$$\mathcal{C}(\ell, t) = e^{-\mu t} (\ell - \gamma e^\ell) + \int_{s \geq t} e^{-\mu s} \left( \mu \chi_s + \gamma e^{\ell + \chi_s} \right) \mathrm{d}s,$$

which is differentiable with respect to $\ell$, with

$$\frac{\partial \mathcal{C}(\ell, t)}{\partial \ell} = e^{-\mu t} \left( 1 - \gamma e^\ell + \gamma e^\ell \int_{s \geq t} e^{\chi_s - \mu(s-t)} \mathrm{d}s \right),$$

which is minimized by setting $\nu_\tau^{-i} = 0$ for all $\tau \geq t$. In that case, the right-hand side is equal to

$$e^{-\mu t} \left( 1 - \gamma e^\ell - \frac{\gamma e^\ell}{\bar{u} + \mu} \right),$$

which is positive if *and only if* $\ell < \ell^*$. Hence, independently of $\nu^{-i}$, $\mathcal{C}(\ell, t)$ is strictly increasing in $\ell$ whenever $\ell < \ell^*$. It follows that, for $\ell < \ell^*$, $\mathcal{C}$ solves the HJB equation

$$\frac{\partial \mathcal{C}(\ell, t)}{\partial t} + \min_{u^i} \left\{ \frac{\partial \mathcal{C}(\ell, t)}{\partial \ell} (u_t^i + \nu_t^{-i} - \bar{u}) + e^{-\mu t} \left( \mu \ell_t - \gamma(\bar{u} - \nu_t^{-i} - 1 + \mu) e^{\ell_t} \right) \right\} = 0,$$

so that setting $u_t^i = 0$ is optimal. Because of the "if and only if" above, if $\nu_s^{-i} = 0$ for all $s \geq t$ (for which it suffices that $\nu_t^{-i} = 0$), yet $\ell_t = \ell > \ell^*$, it cannot be that $u_s^i = 0$ for all $s \geq t$ (and so it must be that $u_t^i > 0$). ■

If $\ell^0 < \ell^*$, we are done: in the unique equilibrium, all players choose $u_t^i = 0$ at all times. In what follows, we implicitly assume $\ell^0 \geq \ell^*$.

**Lemma 4** *If $\nu_t^{-i} = 1$ and $F^i(t) < 1$, then $\ell_t^i \geq \ell^{**}$.*

**Proof.** Ignoring some irrelevant constants, the cost can be rewritten as (abusing notation for $\mathcal{C}$)

$$\mathcal{C}_t^i := \frac{e^{-\mu t}}{\mu} \left( \mu\gamma e^{\ell_t^i} \int_t^\infty e^{\int_t^s (\nu_\tau^{-i} - (\mu+\bar{u}))\mathrm{d}\tau} \mathrm{d}s - 1 \right).$$

If $\nu_t^{-i} = 1$ and $t \in \operatorname{supp} F^i$, yet $\ell_t^i < \ell^{**}$, then because $\mathcal{C}_t^{i'} = 0$,

$$\mathcal{C}_t^{i''} e^{\mu t} = \gamma(\mu + \bar{u} - 2)e^{\ell_t^i} - (\mu + 1) < 0,$$

by definition of $\ell^{**}$. It follows that for small enough $\varepsilon > 0$, $\mathcal{C}_{t-\varepsilon}^i < \mathcal{C}_t^i$, a contradiction. ■

Lemma 4 immediately implies that at least one player must assign positive probability to switching to the risky arm before his belief reaches $\ell^{**}$.

By now, the reader must wonder how $\ell^*$ and $\ell^{**}$ compare. It depends. An immediate computation yields that $\ell^{**} > \ell^*$ if and only if

$$(\bar{u} - I)(\mu + \bar{u}) < \bar{u} - 1. \tag{11}$$

While this inequality can go either way depending on the specific parameters, it is always satisfied in the important case $\bar{u} = I$. Condition (11) plays a key role in the analysis. It immediately follows from Lemma 3–4 that any symmetric equilibrium must be in mixed strategies when 11 holds. Indeed, Lemma 3 implies that at least one (and hence all) players must assign positive probability to switching at the point when their belief reaches $\ell^*$, while Lemma 4 implies that they must also assign positive probability to switching no later than the time at which their belief reaches $\ell^{**}$. Conversely, the unique candidate for a pure strategy equilibrium when (11) does not hold is for players to switch when their belief reaches $\ell^*$. To better understand what is happening, we first examine the structure of best-replies in the case of two players, before solving for the equilibrium.

## 3.4 Best-Replies with Two Players

Here, we examine a player's best-reply when his opponent plays a pure strategy. Suppose player $j \neq i$ switches (with probability one) to the risky arm at time $t^j$. We may distinguish player $i$'s cost according to whether he switches to safe first or second. Accordingly, given some $t^j$, write $\mathcal{C}^F(\cdot; t^j) : [0, t^j] \to \mathbf{R}$ and $\mathcal{C}^S(\cdot; t^j) : [t^j, \infty) \to \mathbf{R}$ for the objective (8) of $\mathcal{P}$, evaluated at a given time $t^i$ at which player $i$ switches to the risky arm.

The proofs of the following claims can be found in Section D of the Appendix.

**Claim 1** *The minimizer $t^i \geq t^j$ of $\mathcal{C}^S(\cdot; t^j)$ is increasing in $t^j$.*

That is, if player $i$ must switch to the risky arm after player $j$, he will do so later, the later player $j$ switches to the risky arm.

Notice that any player who chooses to start experimenting second, will do so when his private belief reaches the threshold $\ell^*$. The implications for the optimal timing are best seen in the case of $\bar{u} = I = 2$. If player $i$ chooses to go second, he will wait until time $t^j + \ell^0 - \ell^*$. The fixed delay is equal to the time required for beliefs to reach the threshold $\ell^*$ based on player $j$'s experimentation alone. With a positive rate of background learning, $\bar{u} > I$, the effect is attenuated, and the slope of the best reply is less than one (see Figure 1), but the intuition is unchanged.

**Claim 2** *If the minimizer of $\mathcal{C}^F(\cdot; t^j)$ is $t^i < t^j$, then it is decreasing in $t^j$ at $t^i$.*

In other words, if player $i$ must switch to the risky arm before player $j$, he will either do it at the same time than player $j$, or if not, he will do so earlier, the later player $j$ switches. Furthermore, if he prefers to switch strictly before player $j$ (rather than at the same time), he also does so for larger values of $t^j$.

Consider again the case of no background learning: conditional on wanting to preempt player $j$, player $i$ will start experimenting immediately. Intuitively, player $i$ will not learn before time $t^j$ unless he experiments. If player $i$ is not willing to wait until then, he should start immediately. If delaying experimentation is not as costly as choosing the risky arm while still pessimistic, then player $i$ will choose to "freeze" beliefs until $t^j$. With a positive rate of background learning, player $i$'s trade-off is smoother. By waiting until $t^i > 0$, he balances the flow cost of playing risky with the value of investing in information. As $t^j$ increases, the marginal value of information increases (*i.e.*, the experimentation levels of the two players are strategic substitutes), and player $i$ chooses to start earlier.

**Claim 3** *There exists $\bar{t}^j \in \mathbf{R}_+$ such that $\min_{t^i \geq t^j} \mathcal{C}^S(t^i; t^j) < \min_{t^i \leq t^j} \mathcal{C}^F(t^i; t^j)$ if and only if $t^j < \bar{t}^j$.*

Hence, player $i$ prefers to switch to the risky arm before player $j$ if and only if player $j$ does so sufficiently late. Figure 1 illustrates the two best-reply curves for values of the parameters such that (11) holds. As a result, the two best-reply curves do not cross and no pure-strategy equilibrium exists. Figure 2 displays these curves for values for which (11) does not hold: the two curves cross, and they do so for a time $t$ such that players' beliefs at that switching time is precisely $\ell^*$.

Recall that condition (11) is always satisfied when $\bar{u} = I$. The previous claims imply that player $i$'s best-reply to $t^j$ can only take the two values 0 and $t^j + \ell^0 - \ell^*$. Intuitively, if $t^j$ is

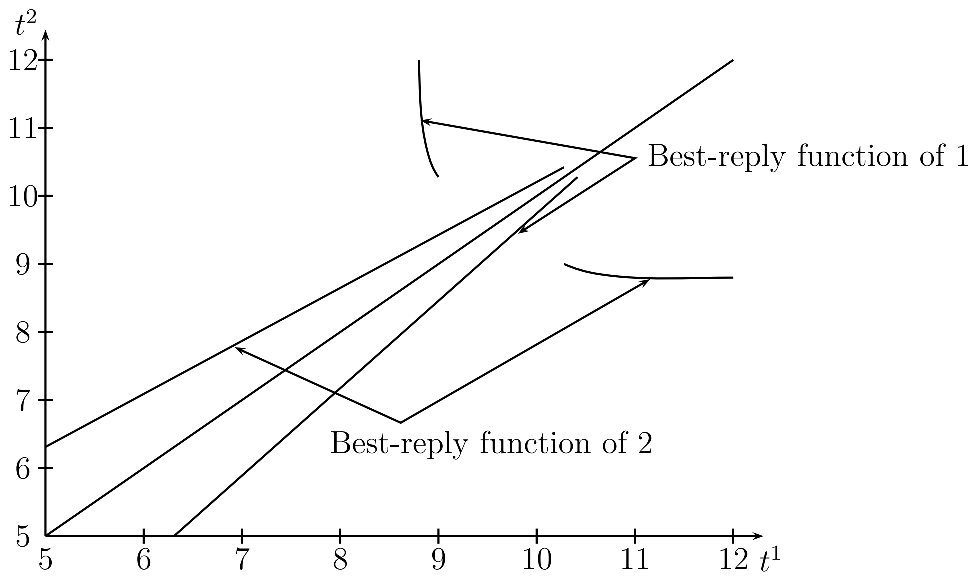



Figure 1: Best-reply curves for $(\ell^0, \mu, \bar{u}, \gamma, I) = (3, 1/5, 23/10, 2, 2)$

very high, then the cost of waiting until player $j$'s actions take beliefs to the threshold causes too costly a delay in learning. Conversely, when player $j$ is expected to switch to risky soon, the benefits of free-riding on his experimentation when beliefs are most pessimistic outweighs the delay cost. Consequently, player $i$'s best reply will have one downward jump.

To gain some intuition for Figure 2, in which the best replies are continuous, consider the opposite case of a very large $\bar{u}$. Loosely speaking, learning is now mostly driven more by background information. In other words, player $i$'s beliefs at time $t$ do not depend as much on whether he is preempting or following player $j$. As a result, a player $j$'s switching time affects player $i$'s best reply continuously.

# 4 Main Results

## 4.1 Symmetric Equilibrium

We now turn to equilibrium analysis. Recall that we assume throughout that $\ell^0 \geq \ell^*$. Given $F^{-i}$, and hence, given $\nu^{-i}$, each time $\tau \in \operatorname{supp} F^i$ must be such that the stopping strategy $\pi^i_\tau$ be a solution to $\mathcal{P}$. Furthermore, it holds that, given any $\tau \in \operatorname{supp} F^i$, $\ell_\tau \geq \ell^*$. We let $\bar{\tau}^i := \max\{\tau \in \mathbf{R}_+ : \tau \in \operatorname{supp} F^i\}$ and $\underline{\tau}^i := \min\{\tau \in \mathbf{R}_+ : \tau \in \operatorname{supp} F^i\}$.

First, we focus on symmetric equilibria, and write accordingly $F$, $\bar{\tau}$ and $\underline{\tau}$ for $F^i$, $\bar{\tau}^i$ and

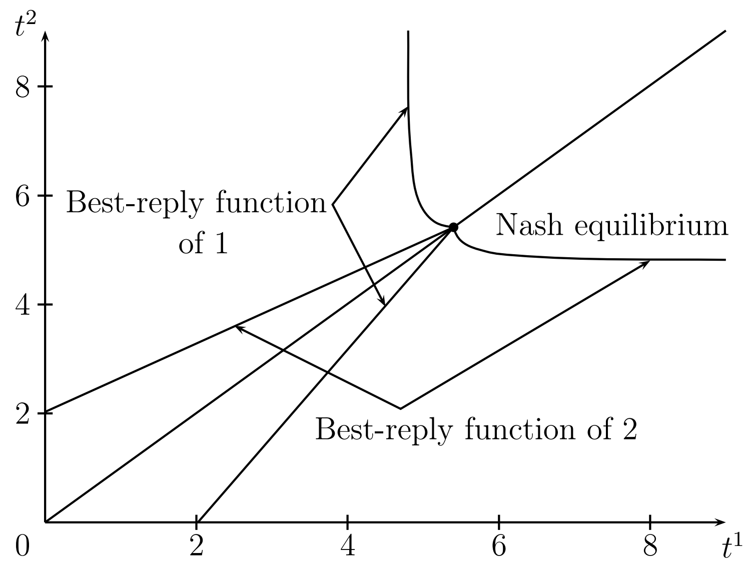



Figure 2: Best-reply curves for $(\ell^0, \mu, \bar{u}, \gamma, I) = (3, 1/5, 26/10, 2, 2)$

$\underline{\tau}^i$, unless we emphasize a given player's perspective. The next result confirms that (11) is a necessary and sufficient condition for the symmetric equilibrium to be in mixed strategies.

**Theorem 4** *There exists a unique symmetric equilibrium.*

1. *If* (11) *holds, this equilibrium involves mixed strategies. Specifically, player $i$ chooses a stopping strategy $\pi_t$ among the set $[\underline{\tau}, \bar{\tau}]$, with $\underline{\tau} < \bar{\tau}$ and $\ell_{\bar{\tau}} = \ell^*$; this distribution is positive and continuous over $(\underline{\tau}, \bar{\tau})$ and has an atom at $\bar{\tau}$. The time $\underline{\tau}$ is strictly positive if and only if $\ell^0 > \ell^{**}$. There is an atom at time $\underline{\tau} = 0$ if and only if $\ell^0 < \ell^{**}$.*

2. *If* (11) *does not hold, this equilibrium involves pure strategies. Specifically, player $i$ chooses $\pi_{t^*}$, where $\ell_{t^*} = \ell^*$.*

The equilibrium distribution function can be solved for in closed-form. See (19)–(20) for the formulas, depending upon whether $\ell^0 \lesseqgtr \ell^{**}$.

**Proof.** We first argue that in every symmetric equilibrium the support of the distribution is an interval: for all $i$, supp $F^i = [\underline{\tau}, \bar{\tau}]$, for some $\underline{\tau} \leq \bar{\tau}$, with $\ell_{\bar{\tau}} = \ell^*$.

Using the same notation as in the proof of Lemma 3, let $\chi_t = \int_{\tau=0}^t (\nu_\tau^{-i} - \bar{u}) \mathrm{d}\tau$. By stopping at time $t$, starting at time 0 with a belief $\ell$, player $i$'s cost is equal to (integrating (8) by parts)

$$\ell - \gamma e^\ell + \int_0^\infty e^{-\mu s} \mu \chi_s \mathrm{d}s + \frac{1 - e^{-\mu t}}{\mu} + \gamma \int_t^\infty e^{\ell + \chi_s - \mu s + t} \mathrm{d}s, \tag{12}$$

which is differentiable in $t$. If $t \in \operatorname{supp} F^i$, it must be that the derivative with respect to $t$ be zero, that is,

$$e^{-\mu t}\left(1-\gamma e^{\ell+\chi_t+t}\right)+\gamma \int_t^{\infty} e^{\ell+\chi_s-\mu s+t} \mathrm{d}s=0. \tag{13}$$

Furthermore, this expression being itself differentiable in $t$, the second derivative must be non-negative, which is equivalent to (differentiating and using the first-order condition)

$$\gamma(\bar{u}-1-\nu_t^{-i}+\mu)-(1+\mu)e^{-\ell_t} \geq 0. \tag{14}$$

Note that the left-hand side of (14) is decreasing in $t$ if $t \notin \cup_{j\neq i} \operatorname{supp} F^j$. Hence, if $t_1, t_2 \in \operatorname{supp} F^i$, with $t_1 < t_2$, it must be that $(t_1, t_2) \cap \operatorname{supp} F^j \neq \emptyset$ for at least one $j \neq i$. Otherwise, (12) must admit a local maximum at some $t \in (t_1, t_2)$, at which value the inequality of (14) is reversed. This is inconsistent with the monotonicity of the left-hand side of (14) over $(t_1, t_2)$, and the fact that it is positive as either $t \downarrow t_1$ or $t \uparrow t_2$. Because we focus on symmetric equilibria, this implies that, for any $t_1, t_2 \in \operatorname{supp} F^i$, $t_1 < t_2$, there exists $t \in (t_1, t_2)$ such that $t \in \operatorname{supp} F^i$. Hence, the support of $F^i$ (a closed set by definition) must be an interval, and by Lemma 3, we must have $\ell_{\bar{\tau}} = \ell^*$.

This leaves us with two cases (the second being further divided depending upon whether $\underline{\tau} = 0$ or $\underline{\tau} > 0$), depending upon whether $\underline{\tau}$ is less than or equal to $\bar{\tau}$. We consider both cases in turn.

1. $\underline{\tau} = \bar{\tau}$: Define $T^*$ as

$$T^* := \frac{\ell^* - \ell^0}{I - \bar{u}}.$$

This is the time it takes for the belief $(\ell_t)$ to reach $\ell^*$ starting at $\ell^0$ under the common strategy that specifies $u_t^i = 0$ if $\ell_t < \ell^*$, $u_t^i = 1$ otherwise. Hence, the unique candidate with $\underline{\tau} = \bar{\tau}$ specifies $\underline{\tau} = \bar{\tau} = T^*$: all players switch from safe to risky at time $T^*$. Given Lemma 2, the only potentially profitable deviations are stopping strategies $\pi_\tau^i$ with $\tau < T^*$. Note that, given that players $j \neq i$ use $\pi_{T^*}^j$, $\nu_t^{-i} = I - 1$ for all $t < T^*$.

Plugging in $t = T^*$ and $\chi_s = (I-1)(s \wedge T^*) - \bar{u}s$ in (13) yields that the (left-)derivative of the cost from stopping at $T^*$ is

$$e^{-\mu T^*}\left(1-\gamma \frac{\mu+\bar{u}-1}{\mu+\bar{u}} e^{\ell^*}\right)=0,$$

given the definition of $\ell^*$. Note that (14) is decreasing in $t < T^*$: the cost from stopping at $t < T^*$ is convex in $t$ if it is convex at $T^*$; hence, a necessary and sufficient condition for $\pi_{T^*}^i$ to be a best-reply to $\{\pi_{T^*}^j : j \neq i\}$ is that this cost be convex at $t = T^*$. Note that the value of (14) at $t = T^*$ is

$$\gamma(\bar{u}+\mu-I)-(1+\mu)e^{-\ell^*}=\gamma\left(\bar{u}+\mu-I-(1+\mu)\frac{\bar{u}+\mu-1}{\bar{u}+\mu}\right), \tag{15}$$

which, using the definition of $\ell^*$, is precisely positive when (11) is violated. Hence, there exists a symmetric equilibrium in which $\underline{\tau} = \bar{\tau}$ if and only if (11) is violated; in this equilibrium, every player uses the stopping strategy $\pi^i_{T^*}$.

2. $0 < \underline{\tau} < \bar{\tau}$ (as we will argue this case cannot occur when $\bar{u} = I$, and we may assume so for the time being): Because the cost from stopping must be constant over $[\underline{\tau}, \bar{\tau}]$, the second derivative given by (14) must be identically zero over $(\underline{\tau}, \bar{\tau})$. Given that $\nu^{-i}_{T^*} \leq I-1$, it follows from the previous inequality (see (15)) that (11) must hold. Eqn. (14) immediately gives $\nu^{-i}_t$ as a function of $\ell_t$. Because $\ell$ is differentiable, so must $\nu^{-i}$ be. Hence, defining $\xi^{-i}_t = (\bar{u} - 1 - \nu^{-i}_t)/\mu$ and differentiating (14) (eliminating $e^{\ell_t}$ by using (14)) gives that $\xi^{-i}$ obeys the differential equation

$$\dot{\xi}^{-i}_t = \mu \xi^{-i}_t (1 + \xi^{-i}_t),$$

and so $\xi^{-i}_t = (A_1 e^{-\mu t} - 1)^{-1}$ for some $A_1 > 0$ (because $\xi^{-i}_t > 0$), yielding

$$\nu^{-i}_t = \bar{u} - 1 + \frac{\mu}{1 - A_1 e^{-\mu t}} \tag{16}$$

for all $t \in (\underline{\tau}, \bar{\tau})$. Hence,

$$\ln \mathbf{E}_{\tau^j}[e^{\int_0^t u^j_s \mathrm{d}s}] = \frac{1}{I-1} \int \left( \bar{u} - 1 + \frac{\mu}{1 - A_1 e^{-\mu s}} \right) \mathrm{d}s = \frac{\ln(A_1 - e^{\mu t})}{I-1} + \frac{\bar{u}-1}{I-1} t + A_2,$$

for some $A_2 \in \mathbf{R}$. That is,

$$\begin{aligned} \int_{s=0}^t e^s \mathrm{d}F(s) + (1 - F(t)) e^t &= e^{A_2} e^{\frac{1}{I-1}\left(\ln(A_1 - e^{\mu t}) + (\bar{u}-1)t\right)} \\ &= e^{A_2} (A_1 - e^{\mu t})^{\frac{1}{I-1}} e^{\frac{\bar{u}-1}{I-1} t}. \end{aligned}$$

Differentiating both sides gives finally

$$1 - F(t) = \frac{e^{A_2}}{I-1} (A_1 - e^{\mu t})^{\frac{1}{I-1}} e^{\frac{\bar{u}-I}{I-1} t} \left( \bar{u} - 1 - \frac{\mu}{A_1 e^{-\mu t} - 1} \right). \tag{17}$$

It remains to determine the constants $A_1, A_2$ as well as $\underline{\tau}$. Assume $\ell^0 > \ell^{**}$. If $\underline{\tau}$ is such that $\ell_{\underline{\tau}} > \ell^{**}$, then solving (14) gives $\nu^{-i}_t > I - 1$ for values of $t$ such that $\ell_t \in (\ell^{**}, \ell_{\underline{\tau}})$ –a contradiction. Hence, $\ell_{\underline{\tau}} \leq \ell^{**}$. If on the other hand $\ell_{\underline{\tau}} < \ell^{**}$, then the left-hand derivative of (14) is strictly negative at $\underline{\tau}$, while by continuity (13) is equal to zero, so that stopping at $\underline{\tau} - \varepsilon$ would yield a strictly lower cost than at $\underline{\tau}$, for $\varepsilon > 0$ small enough, a contradiction as well. Hence $\ell_{\underline{\tau}} = \ell^{**}$ (or $\underline{\tau} = 0$ if $\ell^0 < \ell^{**}$). (The same argument implies at least one $j$ stops no later than $\ell^{**}$ in any equilibrium.) We thus have

$$\underline{\tau} = \frac{\ell^0 - \ell^{**}}{\bar{u} - I}, \tag{18}$$

and so, from (16), we get

$$A_1 e^{-\mu\underline{\tau}} = 1 + \frac{\mu}{\bar{u} - I},$$

yielding

$$A_1 = \left(1 + \frac{\mu}{\bar{u} - I}\right) e^{\mu \frac{\ell^0 - \ell^{**}}{\bar{u} - I}}.$$

Note from (10) that a jump up in $F(t)$ is equivalent to a jump down of $\nu_t^{-i}$. Hence $\nu_{\underline{\tau}}^{-i} = I - 1$ (and so, $\nu^{-i}$ being continuous at $\underline{\tau}$) implies that $F(\underline{\tau}) = 0$. Hence, from (17), we obtain

$$e^{-\frac{\bar{u}-I}{I-1}\underline{\tau}} = (A_1 - e^{\mu\underline{\tau}})^{\frac{1}{I-1}} e^{A_2},$$

Hence,

$$A_2 = \frac{1}{I-1} \ln \frac{(\bar{u} - I)}{\mu} - \frac{\bar{u} + \mu - I}{I - 1} \frac{\ell^0 - \ell^{**}}{\bar{u} - I}.$$

Combining terms, we can write the distribution as

$$\bar{F}(t) = \frac{e^{\frac{\bar{u}-I}{I-1}(t-\underline{\tau})}}{I-1} \left(\bar{u} - 1 - \frac{\mu(\bar{u} - I)}{(\bar{u} + \mu - I)e^{-\mu(t-\underline{\tau})} - (\bar{u} - I)}\right) \left(1 + \frac{(1 - e^{\mu(t-\underline{\tau})})(\bar{u} - I)}{\mu}\right)^{\frac{1}{I-1}}, \tag{19}$$

where $\underline{\tau}$ is given by (18). We have thus characterized a unique candidate symmetric equilibrium. Verification that it is an equilibrium is now immediate: by construction, player $i$ is indifferent over all times $t \in [\underline{\tau}, \bar{\tau}]$; since $\ell_{\bar{\tau}} = \ell^*$, stopping at later times is suboptimal; while stopping at earlier times is also suboptimal, because it follows from (14) that the cost of doing so is convex in $t < \underline{\tau}$, and from (13) that it is decreasing at $t = \underline{\tau}$.

3. $0 = \underline{\tau} < \bar{\tau}$. If $\ell^0 < \ell^{**}$, combine (14) (with equality) at $t = 0$ with (16) to get

$$A_1 = \left(1 - \frac{\mu}{1 + \mu}\gamma e^{\ell^0}\right)^{-1}.$$

Moreover, note from (10) that $1 - F(0) = \nu_0^{-i}/(I-1)$. Plugging in (17) for $t = 0$ using (16) gives $A_2 = (A_1 - 1)^{-\frac{1}{I-1}}$. The resulting distribution is given by

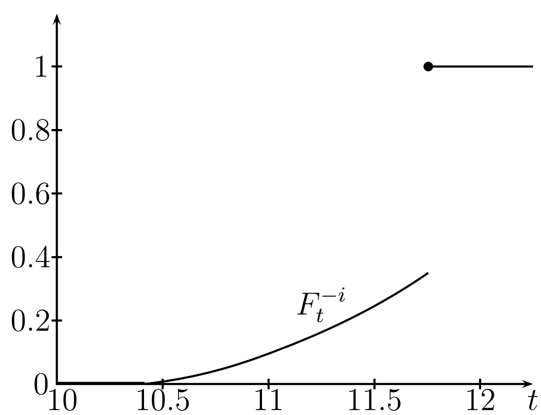
$$\bar{F}(t) = \frac{1}{I-1} e^{\frac{\bar{u}-I+\mu}{I-1}t} \left(\frac{A_1 e^{-\mu t} - 1}{A_1 - 1}\right)^{\frac{1}{I-1}} \left(\bar{u} - 1 - \frac{\mu}{A_1 e^{-\mu t} - 1}\right). \tag{20}$$

The distribution further simplifies to

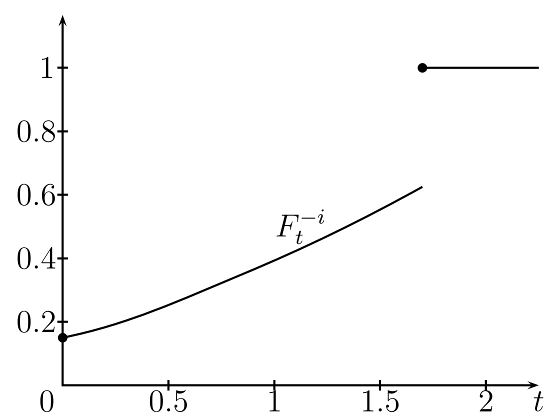
$$\bar{F}(t) = \left(\frac{A_1 - e^{\mu t}}{A_1 - 1}\right)^{\frac{1}{I-1}} \left(1 - \frac{\mu}{(I-1)(A_1 e^{-\mu t} - 1)}\right)$$

in the special case $\bar{u} = I$. Let us make a few final remarks regarding this special case. First, note that this density is 0 at $\ell^0 = \ell^{**}$. That is, if the game starts with this belief, it never changes and the safe arm is used throughout. We must now rule out that $\underline{\tau} > 0$ for this special case. If $\ell^0 = \ell^{**}$, there is nothing to show (as the safe arm is used forever anyhow). If $\ell^0 > \ell^**$, the safe arm must be used throughout (the support of the distribution of stopping beliefs must be convex, yet the cost is strictly quasi-convex in $t$ for $\ell^0 > \ell^**$, yielding a contradiction if this region included a stopping time). If $\ell^0 < \ell^**$, we can use the same argument as in 1. above, based on (14), to conclude that deviating to using the risky arm at time $\underline{\tau} - \varepsilon$ would be a profitable deviation for $\varepsilon > 0$ small enough.

■




(a) $(I, \bar{u}, \mu, \gamma, \ell^0) = (2, 11/5, 1/3, 1, 3)$




(b) $(I, \bar{u}, \mu, \gamma, \ell^0) = (2, 2, 1/3, 1, 1)$

Figure 3: Equilibrium distributions according to whether $\ell^0 \lesseqgtr \ell^{**}$

Figure 3 displays equilibrium distributions for the case in which $\ell^0 \geq \ell^{**}$ (panel (a)), so that there is no atom in the equilibrium distribution at times before the belief reaches $\ell^*$, and for the case in which $\ell^0 < \ell^{**}$ (panel (b)), in which case there is such an atom at time 0.

Figure 4 illustrates the difference between the distribution function $F^{-i}$ and the "contribution" to the hazard rate $\nu^{-i}$, even with two players. As time passes by and no breakdown occurs, players become more optimistic, but the rate at which the usage of the safe arm by the other player slows down this learning process is smaller than the survival rate of the opponent's distribution, because as time passes by a player assigns growing weight to the event in which this opponent has already switched to the risky arm, conditional on which learning should occur faster.

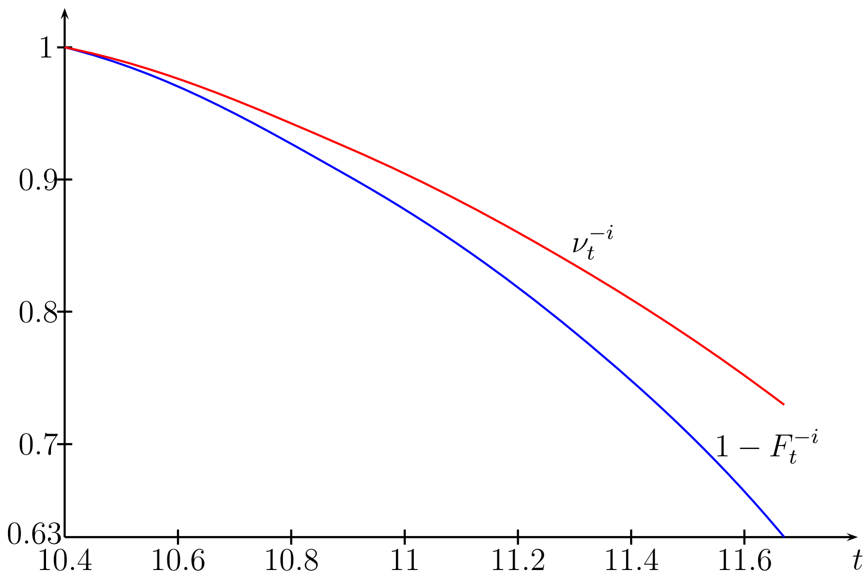



Figure 4: $\nu_t^{-i}$ vs. $F_t^{-i}$, $(I, \bar{u}, \mu, \gamma, \ell^0) = (2, 11/5, 1/3, 1, 3)$

Note that only "on-path" behavior has been described. What is an optimal strategy for player $i$ after some arbitrary history $(u_s^i)_{s=0}^t$ (along which he might have deviated from the prescribed behavior)? First, note that, by Lemma 2–3, it remains the case that player $i$ still has an optimal strategy that is a stopping strategy (from time $t$ on), prescribing stopping no later than the first time at which his belief reaches $\ell^*$. Second, it should also be clear that sufficient statistics for this optimal strategy are his past choices, via $\int_{s=0}^t u_s^i \mathrm{d}s$, and calendar time $t$; this suffices for him to determine his current belief $\ell = \ell_t$ and the future behavior of his opponents, as summarized by $\nu_t^{-i}$. When does stopping occur? Note that (13) rewrites as

$$e^{-(\mu+1)t-\ell} + \gamma \left( \int_t^\infty e^{\chi_s - \mu s} \mathrm{d}s - e^{\chi_t - \mu t} \right) = 0. \tag{21}$$

Hence, if this equation holds on path for $\ell_t = \ell$, then the left-hand side is strictly positive for all lower values of $\ell$, corresponding to player $i$ having experimented more than he was supposed to;[7] hence, if player $i$ has chosen to experiment at times $t \leq \underline{\tau}$ (*i.e.*, $\int_{s=0}^{\underline{\tau}} u_s^i \mathrm{d}s < \underline{\tau}$), it is optimal for him to stop no later than and possibly before $\underline{\tau}$ (or immediately if $t > \underline{\tau}$).

[7] Given the equilibrium strategies, this is the only possible deviation for which optimal behavior is not obvious.

## 4.2 Uniqueness

**Theorem 5** *A pure-strategy equilibrium exists if and only if* (11) *does not hold. If it does, it is unique (and so equal to the symmetric equilibrium).*

**Proof.** We show that every pure strategy equilibrium in the game is symmetric. The strategy is to use induction on the number of players. Because the $I-1$ players who are not the first to stop play the same game (just with a different $\ell^0$), we know by induction hypothesis that they must stop at the same time in equilibrium. We only need to show there does not exist an equilibrium where the first player stops at time $t$ and all other players stop at time $T_2 > T_1$.

Let $\mathcal{U}_1, \mathcal{U}_2$ denote the equilibrium payoff for players 1 and 2. Let $\mathcal{U}_1^*$ denote player 1's payoff should he deviate to stopping at $T_2$. Similarly let $\mathcal{U}_2^*$ denote player 2's payoff from stopping at $T_1$ instead. We show below that

$$\mathcal{U}_1 + \mathcal{U}_2 > \mathcal{U}_1^* + \mathcal{U}_2^*. \tag{22}$$

This inequality would imply that at least one of the deviations is profitable, contradicting equilibrium. To prove (22), we first recall that up to a player-specific constant, each player's cost can be rewritten as

$$\mathcal{C}_t^i := \frac{e^{-\mu t}}{\mu}\left(\mu\gamma e^{\ell_t^i}\int_t^\infty e^{\int_t^s(\nu_\tau^{-i}-(\mu+\bar{u}))\mathrm{d}\tau}\mathrm{d}s - 1\right). \tag{23}$$

It follows that we only need to show

$$\mathcal{C}_{T_1}^1 + \mathcal{C}_{T_2}^2 > \mathcal{C}_{T_1}^2 + \mathcal{C}_{T_2}^1. \tag{24}$$

Plugging (23) into (24), we find that terms $\frac{e^{-\mu t}}{\mu}$ drop out nicely. (24) simplifies to

$$\mathcal{D}_{T_1}^1 + \mathcal{D}_{T_2}^2 > \mathcal{D}_{T_1}^2 + \mathcal{D}_{T_2}^1,$$

where we define

$$\begin{aligned}\mathcal{D}_t^i :&= e^{-\mu t+\ell_t^i}\int_t^\infty e^{\int_t^s(\nu_\tau^{-i}-(\mu+\bar{u}))\mathrm{d}\tau}\mathrm{d}s\\ &= e^{\ell^0}e^t\int_t^\infty e^{\int_0^s(\nu_\tau^{-i}-(\mu+\bar{u}))\mathrm{d}\tau}\mathrm{d}s.\end{aligned}$$

The equality comes from $\ell_t^i = \ell^0 + \int_0^t(u_\tau^i + \nu_\tau^{-i} - \bar{u})\mathrm{d}\tau$ and $u_\tau^i = 1$ for $\tau \in [0,t]$.
We now calculate $\mathcal{D}_{T_1}^1$, $\mathcal{D}_{T_2}^1$, $\mathcal{D}_{T_1}^2$ and $\mathcal{D}_{T_2}^2$ (ignoring the constant $e^{\ell^0}$):

$$\begin{aligned}\mathcal{D}_{T_1}^1 &= e^{T_1}\left(\int_{T_1}^{T_2} e^{(I-1-(\mu+\bar{u}))s}\mathrm{d}s + \int_{T_2}^\infty e^{-(\mu+\bar{u})s+(I-1)T_2}\mathrm{d}s\right)\\ &= e^{T_1}\left(\frac{e^{(I-1-(\mu+\bar{u}))T_1}}{\mu+\bar{u}-I+1} - \frac{(I-1)e^{(I-1-\phi)T_2}}{(\mu+\bar{u}-I+1)(\mu+\bar{u})}\right).\end{aligned}$$

Similarly we obtain

$$\mathcal{D}^1_{T_2} = e^{T_2} \int_{T_2}^{\infty} e^{-(\mu+\bar{u})s+(I-1)T_2} \mathrm{d}s = \frac{e^{(I-(\mu+\bar{u}))T_2}}{\mu+\bar{u}}.$$

We also have

$$\begin{aligned}\mathcal{D}^2_{T_1} &= e^{T_1} \left( \int_{T_1}^{T_2} e^{-(\mu+\bar{u})s+T_1+(I-2)s} \mathrm{d}s + \int_{T_2}^{\infty} e^{-(\mu+\bar{u})s+T_1+(I-2)T_2} \mathrm{d}s \right) \\ &= e^{T_1} \left( \frac{e^{(I-1-(\mu+\bar{u}))T_1}}{\mu+\bar{u}-I+2} - \frac{(I-2)e^{(I-2-(\mu+\bar{u}))T_2+T_1}}{(\mu+\bar{u}-I+2)(\mu+\bar{u})} \right).\end{aligned}$$

Finally

$$\mathcal{D}^2_{T_2} = e^{T_2} \int_{T_2}^{\infty} e^{-(\mu+\bar{u})s+T_1+(I-1)T_2} \mathrm{d}s = \frac{e^{(I-1-(\mu+\bar{u}))T_2+T_1}}{\mu+\bar{u}}.$$

From the above four equations we have

$$\mathcal{D}^1_{T_1} - \mathcal{D}^2_{T_1} = e^{T_1} \left( \frac{e^{(I-1-(\mu+\bar{u}))T_1} - e^{(I-1-(\mu+\bar{u}))T_2}}{(\mu+\bar{u}-I+1)(\mu+\bar{u}-I+2)} - e^{(I-2-(\mu+\bar{u}))T_2} \frac{(I-2)(e^{T_2}-e^{T_1})}{(\mu+\bar{u}-I+2)(\mu+\bar{u})} \right),$$

and it suffices to show that the right-hand side above is greater than

$$\mathcal{D}^1_{T_2} - \mathcal{D}^2_{T_2} = \frac{e^{(I-1-(\mu+\bar{u}))T_2}}{\mu+\bar{u}}(e^{T_2}-e^{T_1}).$$

Dividing by $e^{T_1}e^{(I-1-(\mu+\bar{u}))T_2}$ and letting $x := e^{T_2-T_1} > 1$, the desired inequality becomes

$$\frac{x^{\mu+\bar{u}-I+1}-1}{(\mu+\bar{u}-I+1)(\mu+\bar{u}-I+2)} > \frac{(I-2)(1-\frac{1}{x})}{(\mu+\bar{u}-I+2)(\mu+\bar{u})} + \frac{x-1}{\mu+\bar{u}}. \tag{25}$$

As $x > 1$ and $\mu+\bar{u}-I+1 > 1$, we have

$$\frac{x^{\mu+\bar{u}-I+1}-1}{\mu+\bar{u}-I+1} > x-1,$$

and (25) follows directly, and so does the initial claim. ■

This leaves open the possibility of asymmetric equilibria in mixed strategies, whether (11) holds or not. As Figure 1 makes clear, our game is not supermodular: in particular, best reply curves are not monotone, which makes it difficult to establish uniqueness. The different methods and tricks described in Karlin (1959) do not appear effective either, and indeed we do not know whether the equilibrium is unique in general or not.

**Theorem 6** *If* (11) *does not hold, the unique (pure or mixed) equilibrium of the game is the symmetric pure strategy equilibrium of Theorem 4.*

**Proof.** It follows from (14) that the cost function of player $i$ is strictly quasi-convex in $t$ for $\ell_t^i > \ell^{**}$.[8] Hence, it can have only one minimum for such values. However, Lemma 3 implies that any minimizing time for player must be no less than $\ell^*$, and so, no less than $\ell^{**}$ whenever (11) does not hold. Hence, any equilibrium must be in pure strategies. The result follows from Theorem 5. ■

This only leaves open the problem of uniqueness in mixed strategies when (11) holds. Could there be another, necessarily mixed, necessarily asymmetric equilibrium in that case? We do not know, but conjecture that the answer is negative, at least with two players.

**Conjecture 7** *Assume that $I = 2$. The equilibrium is unique (and so equal to the mixed-strategy equilibrium of Theorem 4 whenever* (11) *holds).*

**Big missing step: prove conjecture, at least for $\bar{U} = I$.**

# 5 Comparative Statics and Benchmarks

## 5.1 Comparative Statics

We consider how equilibrium actions and costs are affected by the parameters of the model, and in particular by the number of players $I$, by the normalized discount rate $\mu$, and by the relative cost of a breakdown $\gamma$. For the case of unobservable actions, we focus on parameter values that yield a mixed-strategy equilibrium, *i.e.* that (11) holds. Finally, to perform comparative statics with respect to $I$, we fix a (strictly positive) level of background learning $\bar{u} - I$, *i.e.*, we let the number of players and the baseline arrival rate of a breakdown $\bar{u}$ increase at the same rate.

We begin by considering the timing of experimentation in all three cases, as the number of players increases. For the unobservable case, we take the perspective of an outside observer. We denote the first time at which $\nu_t < I$ as the beginning of experimentation, and the first time at which $\nu_t = 0$ as the beginning of full experimentation.

**Proposition 1**

1. *In the unobservable case:*

    (a) *experimentation begins at time $\underline{\tau}$, which is constant in $I$;*

[8] One could add to Lemma 4 the observation that a player can only assign positive probability to switching to the risky arm at one time $t$ among those times $t$ for which $\ell_t^i > \ell^{**}$.

(b) *full experimentation begins at time* $\bar{\tau}_I$*, which is increasing in* $I$*;*

(c) *the stopping-time distributions* $F_t^I$ *are ranked by first-order stochastic dominance;*

(d) *for an outside observer,* $\nu_t^{I'} \geq \nu_t^I$ *for all* $I' > I$*, with strict equality for* $t < \bar{\tau}_I'$*;*

(e) *the belief path* $p_t^I$ *crosses* $p_t^{I'}$ *once (from above).*

2. *In the observable case, experimentation begins earlier as* $I$ *increases.*

3. *In the first-best case, the belief paths are ranked for all* $t$*, and* $p_t^{I'} \leq p_t^I$ *for all* $I' > I$*, with strict inequality for* $t > t_{I'}^{FB}$*.*

**Proof.** (1.) In the unobservable case, we establish the following results.

The first threshold $\underline{\tau}$ is equal to

$$\frac{1}{\beta}\left(\ell^0 - \ln\left(\frac{1+\mu}{\gamma(\beta+\mu)}\right)\right).$$

It is then constant in $I$. The second threshold $\bar{\tau}_I$ satisfies

$$e^{-\mu(\bar{\tau}-\underline{\tau})} = \frac{1+\frac{\mu(\mu+\bar{u})}{\bar{u}-1}}{1+\frac{\mu}{\bar{u}-I}}.$$

This follows from (11), (14) and $\ell_{\bar{\tau}} = \ell^*$. Because the RHS decreases in $I$, it follows that $\bar{\tau}$ increases in $I$.

The stopping-time distribution (when $0 < \underline{\tau} < \bar{\tau}$) is given by

$$F(t)^I = 1 - \left(e^{(\beta+\mu)(t-\underline{\tau})}\frac{(\beta+\mu)e^{\mu(\underline{\tau}-t)}-\beta}{\mu}\right)^{\frac{1}{I-1}}\left(1+\beta\frac{1-\frac{\mu}{(\beta+\mu)e^{\mu(\underline{\tau}-t)}-\beta}}{I-1}\right).$$

For a given $t > \underline{\tau}$, the first term is smaller than one, hence increasing in $I$, and the second term is increasing in $I$. Therefore, the partial derivative of $F_t^I$ with respect to $I$ is positive. In addition, $F(t)^I = 0$ for all $I$ and $t \leq \underline{\tau}$, while the upper bound of the support increases with $I$, so that the atom occurs later. Therefore, $F(t)^{I'} \leq F_t^I$ for all $I' > I$ and for all $t$.

From the outside observer's perspective,

$$-\dot{\ell}_t^I = \beta + I - \nu_t^I.$$

Aggregating the individual actions $\nu_t^i$, we obtain

$$-\dot{\ell}_t^I = \beta\left(\frac{I}{I-1}\left(\frac{\mu}{(\beta+\mu)e^{\mu(\tau-t)}-\beta}-1\right)+1\right),$$

which is decreasing in $I$ as the first term in parentheses is positive. Therefore, when $t \in [\underline{\tau}, \bar{\tau}]$, the hazard rate of a breakdown conditional on the bad state is decreasing in $I$. This implies $\nu_t^I$ is increasing in $I$, since the hazard rate is increasing in $I$ for a fixed $\nu$.

Finally, because experimentation begins at a constant time $\underline{\tau}$, and because the derivative of beliefs $\dot{\ell}_t$ is monotone in $I$ over the mixing region, it follows that the beliefs at the end of this phase are ordered $\ell^I_{\bar{\tau}_I} < \ell^{I'}_{\bar{\tau}_{I'}}$ for $I' > I$. (Note that the outside observer's beliefs at $\bar{\tau}$ are more optimistic than $\ell^*$.) However, because the arrival rate of a breakdown is given by $\bar{u}$ when experimentation is full, the observer's beliefs for $I' > I$ will eventually catch up. Hence the two belief paths coincide until $\underline{\tau}$ and cross once thereafter.

(2.) In the observable case, experimentation begins at a belief threshold $\bar{\ell}_I$ which can be written as

$$\bar{\ell}_I =: \ell^*_I - 1 - \frac{\beta+1}{\mu} - W_{-1}\left(-\gamma \frac{\beta+\mu}{\mu} e^{\ell^*-1-\frac{\beta+1}{\mu}}\right).$$

Therefore, $\bar{\ell}_I$ depends on $I$ only through $\ell^*$, which is decreasing in $I$. Finally, notice that

$$\frac{\partial \bar{\ell}}{\partial \ell^*} = \left(1 + W_{-1}\left(-\gamma \frac{\beta+\mu}{\mu} e^{\ell^*-1-\frac{\beta+1}{\mu}}\right)\right)^{-1} < 0,$$

which means experimentation begins earlier (at a higher $\bar{\ell}$) as $I$ increases.

(3.) In the first-best case, the threshold for full experimentation $\tau^{FB}$ is decreasing in $I$. Hence belief paths for $I'$ and $I < I'$ coincide until the earlier threshold $\tau^{FB}_{I'}$ and are ranked for all higher $t$. ■

To summarize the results for the unobservable case: as the number of players increases, the "mixing phase" (which always begins at the same time) lasts longer (until $\bar{\tau}_I$) and drives beliefs to a lower threshold $\ell^*_I$. In particular, the distributions of stopping times $F_t$ with different $I$ are ranked by first-order stochastic dominance: a higher number of players makes later times more likely. Furthermore, the expected hazard rate from the outside observer's perspective $\bar{u} - \nu_t^I$ is decreasing in $I$ as long as $\ell^*$ has not been reached. More precisely, the hazard rate for $I' > I$ jumps above the hazard rate for $I$ as soon as $\nu$ jumps to zero. The outside observer's beliefs $I' > I$ will then eventually overtake the beliefs for $I$.

In Figure 5 we illustrate the hazard rate $\bar{u} - \nu_t^I$ and the belief paths for $I = 2, 4$.

Finally, in the unobservable case, the magnitude of the lump-sum losses affects the timing of experimentation only: as the relative cost of a breakdown $\gamma$ increases, the thresholds $(\underline{\tau}, \bar{\tau})$ shift forward. In other words, a higher $\gamma$ delays the beginning of experimentation and the beginning of full experimentation by an equal amount, without affecting the equilibrium distribution of stopping times.

We now turn to comparative statics of equilibrium costs. We define the cost as the insurable component of the original problem, *i.e.*, the equilibrium value of the objective in

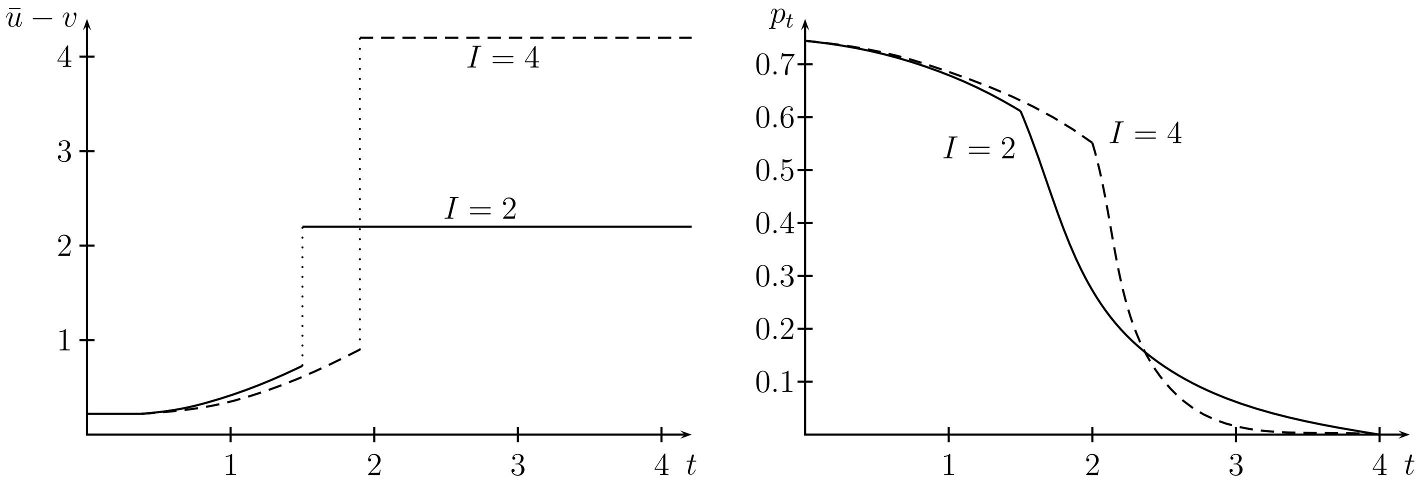



Figure 5: Hazard Rates and Belief Paths, Unobservable Case $(\bar{u},\mu,\gamma,\ell^0)=(I+1/4,1/4,1,1)$

problem $\mathcal{P}$.

**Proposition 2**

1. *The cost $\mathcal{C}(p^0)$ in the first-best and in the symmetric Markov equilibrium of the observable case is strictly decreasing in $I$.*

2. *As $I\to\infty$, the first-best cost converges to $p^0 s$. In the observable case, if $\ell^0>-\ln\gamma$, the cost converges to a limit strictly above $p^0 s$.*

3. *In the unobservable case, the symmetric equilibrium cost is constant in $I$.*

4. *In all three cases, costs converge to $p^0 s$ as $\mu\to 0$.*

**Proof.**

1. Let $\beta=\bar{u}-I$. In the first-best case, we can use the cost formulation in (28) and write

$$\begin{aligned}\mathcal{C}^{FB}(\ell^0) &= \frac{s(1+\mu\gamma)}{1+e^{-\ell^0}}+\frac{\mu s}{1+e^{\ell^0}}\int_0^{\tau^{FB}} e^{-\mu t}\left(-\beta\mu t-\gamma(\beta+\mu)e^{\ell^0-\beta t}+1+\beta\right)\mathrm{d}t\\ &\quad+\frac{\mu s}{1+e^{\ell^0}}\int_{\tau^{FB}}^{\infty} e^{-\mu t}\left(\mu(-\bar{u}t+I\tau^{FB})-\gamma(\bar{u}-1+\mu)e^{\ell^0-\bar{u}t+I\tau^{FB}}+\bar{u}\right)\mathrm{d}t,\end{aligned}$$

where $\tau^{FB}=(\ell^0-\ell^{FB})/\beta$ and $\ell^{FB}=\ln(\beta+I+\mu)-\ln(\gamma(\beta+\mu))$. Integrating and simplifying yields

$$\frac{\mathcal{C}^{FB}(\ell^0)}{s}=1-\frac{\beta}{\beta+\mu}\frac{e^{-\mu\tau^{FB}}}{1+e^{\ell^0}},$$

which is decreasing in $I$ since $\tau^{FB}$ is decreasing in $I$.

In the observable case, we know from the proof of Theorem 8 that the cost over the mixing region must be constant. With our current formulation of the objective, the cost must be equal to $\ell^* - 1$. Over the first phase then, the cost $\mathcal{C}(\ell)$ solves the following boundary value problem:

$$\mu \mathcal{C}(\ell) = \mu \ell - \gamma(\mu + \beta)e^{\ell} + 1 + \beta - \beta \mathcal{C}'(\ell),$$

subject to

$$\mathcal{C}(\bar{\ell}) = \ell^* - 1.$$

Using the cost formulation (28), equilibrium cost is given by

$$\begin{aligned} \frac{\mathcal{C}^o(\ell^0)}{s} &= \frac{1}{1+e^{\ell^0}} \left( \mu \mathcal{C}(\ell^0) + (1+\mu\gamma)e^{\ell^0} - \mu \ell^0 \right) \\ &= 1 - \frac{e^{\frac{\mu(\bar{\ell}-\ell^0)}{\beta}} \left( \mu \left( -\gamma e^{\bar{\ell}} + \bar{\ell} - \ell^* + 1 \right) + 1 \right)}{e^{\ell^0}+1}, \end{aligned}$$

and differentiating with respect to $I$ (which only enters through the threshold beliefs) yields

$$\frac{\mathrm{d}\mathcal{C}^o}{\mathrm{d}I} = -\frac{e^{-\frac{\mu(\bar{\ell}-\ell^0)}{\beta}} \mu}{\left(e^{\ell^0}+1\right)(\beta + I + \mu - 1)(\beta + I + \mu)}.$$

2. Given the prior belief $\ell^0$, for sufficiently high $I$, the first-best policy consists of full experimentation from time 0 on. The cost is then given by

$$\frac{\mathcal{C}^{FB}(\ell^0)}{s} = 1 - \frac{1}{e^{\ell^0}+1} \left( 1 - \frac{\gamma e^{\ell^0} \mu}{\beta + I + \mu} \right),$$

which converges to $s(1+e^{-\ell^0})^{-1} = p^0 s$ as $I \to \infty$.

In the observable case, it suffices to show that $\ell^*$ converges to $-\ln \gamma$ and that $\bar{\ell}$ depends on $I$ only through $\ell^*$. Moreover, because $\mathcal{C}^o(\ell^0)$ depends on $I$ only through these two thresholds, costs also converge. Finally, they must converge to a value strictly above $sp^0$ because $\bar{\ell}$ is finite while $\ell^{FB}$ grows without bound.

3. In the unobservable case, we consider the cost of the most pessimistic type, *i.e.*, of the player who chooses $u_t^i = 1$ until $\bar{\tau}$. Using cost formulation (28) and solving for the equilibrium beliefs from the indifference condition (14), we can write

$$\begin{aligned} (\mathcal{C}^n(\ell^0) - \frac{s(1+\mu\gamma)}{1+e^{-\ell^0}}) \frac{1+e^{\ell^0}}{\mu s} &= \int_0^{\underline{\tau}} e^{-\mu t} \left( -\beta \mu t - \gamma(\beta+\mu) e^{\ell^0 - \beta t} + 1 + \beta \right) \mathrm{d}t \\ &+ \int_{\underline{\tau}}^{\bar{\tau}} e^{-\mu t} \left( \mu \left( \ln \frac{1+\mu}{\gamma(\bar{u}+\mu-1-\nu_t)} - \ell^0 \right) - 1 - \mu + \bar{u} - \nu_t \right) \mathrm{d}t \\ &+ \int_{\bar{\tau}}^{\infty} e^{-\mu t} \left( \mu \left( \ell^* - \bar{u}(t-\bar{\tau}) - \ell^0 \right) - \gamma(\bar{u} - 1 + \mu) e^{\ell^* - \bar{u}(t-\bar{\tau})} + \bar{u} \right) \mathrm{d}t, \end{aligned}$$

where

$$\nu_t = I - 1 + \beta\left(1 - \frac{\mu}{e^{-\mu(t-\underline{\tau})}(\beta+\mu) - \beta}\right).$$

Simplifying yields

$$\frac{\mathcal{C}^o(\ell^0)}{s} = 1 - \frac{1}{1+e^{\ell^0}}e^{-\mu\underline{\tau}}\frac{\beta}{\beta+\mu}(1+\mu),$$

which is constant in $I$.

4. The costs in all three cases are continuous in $\mu$, and substituting $\mu = 0$ yields the desired result, $\mathcal{C} \to p^0 s$.

■

In other words, lack of monitoring implies full dissipation of the positive informational externalities generated by an additional player. We should also add that the cost (cost) in the case of a pure-strategy equilibrium is decreasing in the number of players. However, note that as $I$ increases we can only move from pure- to mixed-strategy equilibrium, with costs "leveling off" for sufficiently large $I$.

Under observable actions, players benefit from a larger group, though not at the same rate as the social planner. Furthermore, as the number of players (and hence the value of information) grows without bound, the first-best policy consists of immediate full experimentation. Each agent's cost then converges to the complete-information level: the arrival rate of a breakdown, conditional on the bad state, grows without bound and the probability of suffering a breakdown is inversely proportional to $I$. This cannot be the case when actions are not observable, because experimentation does not even begin until time $\underline{\tau}$. Even under observable actions, the threshold belief $\bar{\ell}_I$ for beginning experimentation is increasing in $I$, but converges to a finite value. Furthermore, the duration of the mixing phase does not vanish as the number of players grows. Therefore, the complete-information cost is not attainable if $\ell^0 > \ell^*$.

Note that $\beta > 0$ implies complete learning in our model. Thus, in the undiscounted limit, players attain the complete-information expected cost, which is given by $p^0 \cdot s + (1 - p^0) \cdot 0$.

## 5.2 Observable Actions

Here we recall (and very slightly extend) Keller and Rady's result regarding symmetric Markov equilibria in the game with observable actions. Players are restricted to Markov strategies $u^i : \mathbf{R} \to [0, 1]$ with the left limit $\ell_{t-}$ of the common posterior belief as state variable. Strategies are required to be left-continuous and piecewise Lipschitz. Let

$$\bar{\ell} := \ell^* - 1 - \frac{\bar{u} - I + 1}{\mu} - W_{-1}\left(-\gamma\frac{\bar{u} - I + \mu}{\mu}e^{\ell^* - 1 - \frac{\bar{u}-I+1}{\mu}}\right),$$

where $W_{-1}$ is the (negative branch of the) Lambert function. We also write $\bar{p}$ for the corresponding probability $\bar{\ell}/(1+\bar{\ell})$. Define also $u^o : \mathbf{R} \to [0,1]$ as

$$u^o(\ell) = \begin{cases} 1 & \text{if } \ell \geq \bar{\ell}, \\ \frac{\bar{u}+\mu-1}{I-1} - \frac{\mu(\ell-\ell^*)+1}{(I-1)(\gamma e^\ell - 1)} & \text{if } \ell \in [\ell^*, \bar{\ell}), \\ 0 & \text{if } \ell < \ell^*. \end{cases}$$

**Theorem 8 (Keller & Rady, 2014)** *The unique symmetric Markov equilibrium is given by $u^o$.*

It is worth emphasizing that this is not the unique Markov equilibrium: asymmetric Markov equilibria exist, and the ranking in terms of welfare can go either way. (See Section 3.3 of Keller and Rady, 2014.)

Because Theorem 8 is essentially a result established by Keller and Rady (2014), we only sketch the argument. Because players $-i$ use a Markov strategy, the function $\nu^{-i}$ is now a function of the belief $\ell$. Recall that player $i$ minimizes

$$\int_{t\geq 0} e^{-\mu t} \left( \frac{\mu g\,(1-u_t^i)}{1+e^{-\ell_t}} + \mu u_t^i s + \frac{(\bar{u} - u_t^i - \nu^{-i}(\ell_t))s}{1+e^{-\ell_t}} \right) \frac{1+e^{\ell_t}}{1+e^{\ell_0}} \mathrm{d}t,$$

subject to

$$\dot{\ell}_t = u_t^i + \nu^{-i}(\ell_t) - \bar{u}.$$

This is an autonomous problem, with state variable $\ell$. The Hamilton-Jacobi-Bellman equation for player $i$ is[9]

$$\mu\mathcal{C}(\ell) = \mu\ell - \nu^{-i}(\ell) - \gamma(\bar{u} - \nu^{-i}(\ell) - 1 + \mu)e^\ell + \min_{u^i} \left\{ (u^i + \nu^{-i}(\ell) - \bar{u}) \frac{\partial \mathcal{C}(\ell)}{\partial \ell} \right\},$$

so that $u^i \in \{0,1\}$ unless $\mathcal{C}(\ell) = \text{Cst.} =: \mathcal{C}_1$. Naturally, we expect two thresholds, $\ell^* \leq \bar{\ell}$, such that $u^i = 0$ (use the risky arm) for $\ell \leq \ell^*$, $u^i$ is interior on $[\ell^*, \bar{\ell}]$, and $u^i = 1$ above. When $u^i = 0$, the value is

$$\begin{aligned} \mathcal{C}(\ell) &= \int_{t\geq 0} e^{-\mu t} \left( \mu(\ell - \bar{u}t) - \gamma(\bar{u} - 1 + \mu)e^{\ell - \bar{u}t} \right) \mathrm{d}t \\ &= \ell - \frac{\bar{u}}{\mu} - \frac{e^\ell(\mu + \bar{u} - 1)\gamma}{\mu + \bar{u}}, \end{aligned}$$

so that in this range

$$\frac{\partial \mathcal{C}(\ell)}{\partial \ell} = 1 - \frac{e^\ell \gamma(\mu + \bar{u} - 1)}{\mu + \bar{u}},$$

[9] Its validity is confirmed *ex post* by a standard verification theorem, see Fleming and Soner (2005).

which is positive if and only if $\ell \leq \ell^*$. So

$$\mathcal{C}_1 = \mathcal{C}(\ell^*) = \ell^* - \frac{\bar{u}}{\mu} - 1.$$

When $\bar{u}$ is interior, we have

$$\mu\mathcal{C}_1 = \mu\ell - (I-1)u^i(\ell) - \gamma(\bar{u} - (I-1)u^i(\ell) - 1 + \mu)e^\ell,$$

which can be solved for $u^i$, so that our candidate for $\bar{\ell}$ is the value of $\ell$ for which $u^i(\ell) = 1$, namely the solution to

$$\mu\left(\ell^* - \frac{\bar{u}}{\mu} - 1\right) = \mu\ell - (I-1) - \gamma(\bar{u} - I + \mu)e^\ell,$$

whose solution is indeed $\bar{\ell}$, and indeed $u^i = u^o$.

Next, we would like to compare the total amount of experimentation up to some $t$ under both observable and unobservable actions. We are faced with a difficulty here, as the unobservable strategies are a function of time, while the observable strategies are a function of the belief. Furthermore, in the unobservable case, the belief that player $i$ holds at a given time is not uniquely pinned down in the mixed equilibrium: the earlier a player stops, the lower his belief at a given time $t$, provided no breakdown occurred.

We are thus led to take the point of view of an *outside observer* who observes nothing at all: conditional on a given time $t$ being reached without a breakdown under either informational assumption, what probability does he attach to $\omega = B$? In the observable case, this belief coincides with that of any of the players, at least on path. In the unobservable case, it is some weighted average of a player's belief, where the weight reflects the probability attached by this observer that a player stops playing safe at a given time, suitably updated given that time $t$ is reached without a breakdown. Formally, we compute

$$p_t = \mathbf{P}^\phi_{p^0}[\omega = B \mid \varnothing_t],$$

where, unlike in Section 3.1 we do not condition on any particular player's action path. It follows that the outside observer's belief satisfies

$$\dot{\ell}_t = I\nu^i_t - \bar{u}.$$

We write $p^o_t$, $p^n_t$ and $p^{FB}_t$ for these beliefs, depending on whether we consider the observable, unobservable or cooperative case. The comparison (in log-likelihood ratio terms, with the obvious notation), is illustrated in Figure 5.2

As it turns out, we can prove a stronger result than the ranking of the belief paths. For the unobservable case, let $\nu(p) := \nu_{t(p)}$, where $t(p)$ denotes the time at which the outside observer's belief reaches a value of $p$. Formally, we can show that $\nu(p)$ is ranked across the three cases.

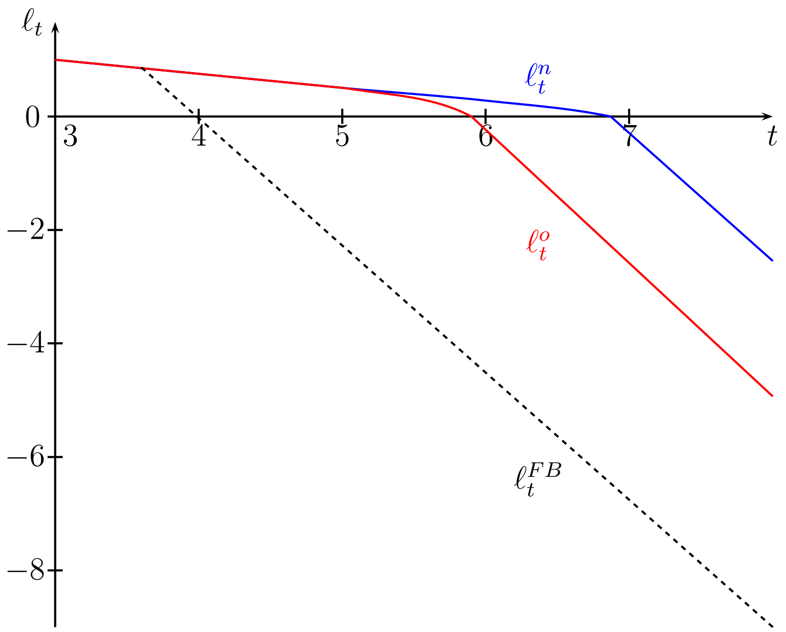



Figure 6: Observable, unobservable and first-best beliefs, $(I, \bar{u}, \mu, \gamma, \ell^0) = (2, 23/10, 1/5, 2, 2)$

**Proposition 3** *The following inequalities hold for all $p$:*

$$\nu^n(p) \geq \nu^o(p) \geq \nu^{FB}(p).$$

*The second inequality is strict whenever $p < p^{FB}$ and the first is strict whenever $p^o < \bar{p}$. In particular, $\bar{p} > p^{**}$.*

Notice that this result implies the ranking of the belief paths shown in Figure 5.2. Therefore, for all $t$,

$$p_t^n \geq p_t^o \geq p_t^{FB},$$

with strict inequalities as described in the previous proposition. Furthermore, Lemma 1 implies the ranking of the symmetric equilibrium costs $\mathcal{C}(p^0)$.

**Corollary 9** *The following inequalities hold for all $p = p^0$:*

$$\mathcal{C}^n(p) \geq \mathcal{C}^o(p) \geq \mathcal{C}^{FB}(p).$$

*Both inequalities are strict whenever $p > p^*$ and $\mu > 0$.*

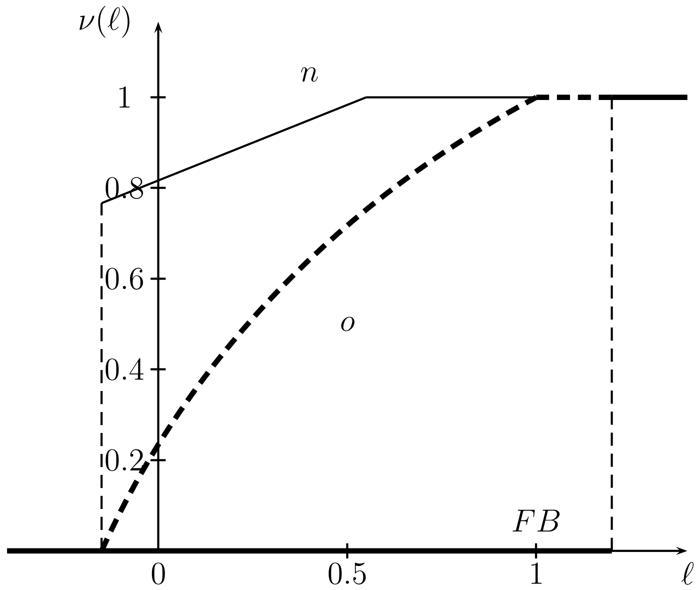



Figure 7: Observable, unobservable and first-best beliefs, $(I, \bar{u}, \mu, \gamma, \ell^0) = (2, 23/10, 1/5, 2, 2)$

### 5.3 Asymmetric Players

One might wonder to what extent our results rely on the symmetry among players. As Figure 1 makes clear, as long as best-reply curves vary continuously with parameters, the non-existence of pure-strategy equilibrium when (11) holds is robust to slight perturbations in parameters. It is legitimate nonetheless to ask how the equilibrium looks like when players have distinct cost functions, whether the equilibrium is mixed or not (and what condition pins down the existence of an equilibrium in pure strategies).

**Details to be completed**

## 6 Discussion

The comparison between our results and the existing literature is surprising in several respects.

First, in comparison with Bonatti and Hörner (2011) who consider unobservable actions with good news, we obtain equilibrium uniqueness with bad news, while there multiple equilibria exist with good news. The contrast is all the more surprising in light of the literature on repeated games with imperfect public monitoring (unobservable actions being a special case) suggesting that multiplicity is more likely to arise with bad news instead of good news (see Abreu, Milgrom and Pearce, 1991). This overlooks some important differences. First, in Bonatti and Hörner (2011), there are not only informational externalities, but also payoff externalities: a breakthrough is shared equally among players. Clearly, payoff externalities make it easier to sustain more equilibrium outcomes, as it increases the scope for actions of others to discipline a player's behavior. Second, the cost function in Bonatti and Hörner (2011) is linear, and this plays an important role in the construction of asymmetric equilibria. With strictly convex costs (with zero derivative at zero effort), it is no longer an equilibrium for a player not to put in any effort, resulting in effort equalization across players, and we conjecture that the equilibrium would be unique with convex costs. In contrast, if we assumed cost externalities in the present paper, multiple equilibria could be constructed in some cases.[10]

For instance, if a breakdown entails the same cost to all players independently of the identity of the player who caused it, and the state of the world was known to be bad, then one could use constructions as in Abreu, Milgrom and Pearce (1991) to construct multiple, Pareto-ranked equilibria. This is the case even if the identity of the player who caused the breakdown remains unknown (unlike in the model discussed at the end of this section, but as in Abreu, Milgrom and Pearce, 1991). While this might be standard to some, we believe that it is valuable to substantiate this claim. Consider the following strategy profile, implementable by a two-state automaton, with costs $\bar{\mathcal{C}} > \underline{\mathcal{C}}$ (used to denote states as well), such that players must play the safe arm in the state $\underline{\mathcal{C}}$, and the risky arm in the other state. Assuming (for convenience) a public randomization device, suppose that the strategy switches from one state to the other with probability $\underline{\mu}$ (resp., $\bar{\mu}$) from state $\underline{\mathcal{C}}$ to state $\bar{\mathcal{C}}$ (resp., from state $\bar{\mathcal{C}}$ to state $\underline{\mathcal{C}}$), while the state does not change in the absence of a breakdown.[11] Writing down the Bellman equations that correspond to optimality of such behavior given that other players follow it, we must have, in state $\underline{\mathcal{C}}$, to the first order,

$$\begin{aligned}\underline{\mathcal{C}} &= rs\mathrm{d}t + \lambda(\bar{u} - I)\mathrm{d}t(rg + \underline{\mu}\bar{\mathcal{C}} + (1-\underline{\mu})\underline{\mathcal{C}}) + (1 - r\mathrm{d}t - \lambda(\bar{u}-I)\mathrm{d}t)\underline{\mathcal{C}} \\ &\leq \lambda(\bar{u} - I + 1)\mathrm{d}t(rg + \underline{\mu}\bar{\mathcal{C}} + (1-\underline{\mu})\underline{\mathcal{C}}) + (1 - r\mathrm{d}t - \lambda(\bar{u}-I+1)\mathrm{d}t)\underline{\mathcal{C}},\end{aligned}$$

and in state $\bar{\mathcal{C}}$,

$$\begin{aligned}\bar{\mathcal{C}} &= \lambda\bar{u}\mathrm{d}t(rg + \bar{\mu}\underline{\mathcal{C}} + (1-\bar{\mu})\bar{\mathcal{C}}) + (1 - r\mathrm{d}t - \lambda\bar{u}\mathrm{d}t)\bar{\mathcal{C}} \\ &\leq rs\mathrm{d}t + \lambda(\bar{u}-1)\mathrm{d}t(rg + \bar{\mu}\underline{\mathcal{C}} + (1-\bar{\mu})\bar{\mathcal{C}}) + (1 - r\mathrm{d}t - \lambda(\bar{u}-1)\mathrm{d}t)\bar{\mathcal{C}}.\end{aligned}$$

It is then straightforward to maximize $\bar{\mathcal{C}}$ subject to these constraints, over $\bar{\mu}, \underline{\mu} \in [0,1]$. This

[10]In particular, the absence of cost externality is key to the argument that behavior after the first breakdown, which reveals the state of the world, is trivial, as playing safe is strictly dominant.

[11]Proving the optimality of this structure of transitions is standard and omitted.

gives

$$\bar{\mu} = 1, \underline{\mu} = 0,$$

as well as

$$\underline{\mathcal{C}} = (1 + (\bar{u} - I)(1+\gamma)\lambda)\,\frac{rs}{1+r}, \quad \bar{\mathcal{C}} = \frac{1 + (1+r)(1+\gamma) + (\bar{u} - I)(1+\gamma)\lambda}{1 + r + \lambda \bar{u}}\,\frac{\bar{u}\lambda rs}{1+r},$$

(with indeed $\underline{\mathcal{C}} < \bar{\mathcal{C}}$), as well as one restriction on the parameters, namely,

$$(1 + r - \bar{u})(1+\gamma) \leq \frac{1+r}{\lambda} + \frac{(I-1)(1+r) + \bar{u}}{1 + r + (\bar{u} - I)\lambda},$$

automatically satisfied when $1 + r < \bar{u}$.[12]

To conclude, with bad news, when payoff externalities are strong, multiple equilibria exist when players are patient enough, even when neither effort nor the identity of the player suffering the breakdown is publicly observed.

Of course, when actions are observable, multiple equilibria exist, even within the class of Markov equilibria, and even without payoff externalities, whether news is good or bad. See Keller, Rady and Cripps (2005) for the good news case, and Keller and Rady (2014) for the bad news case.

Second, observability turns out to improve welfare with bad news, while it is detrimental with good news. (The latter result, from Bonatti and Hörner, 2011 does not rely on the payoff externality that they assume.) This contrast is easy to understand. Using the safe arm slows down learning in both models. But slowed learning results in more experimentation by others with good news (slower learning means more optimism and so more experimentation), while it results in less experimentation with bad news (slower learning means more pessimism and so less experimentation). More experimentation by others is always desirable in terms of the informational externality, whether news is good or bad, as it results in more learning. With breakthroughs, more experimentation by others is also desirable in terms of payoff externalities. Hence, because an observable deviation to the safe arm prompts more experimentation by others with good news, observability exacerbates free-riding and under-experimentation in the good news case (under-experimentation being also present with unobservable actions). With bad news, more experimentation by others is *not* desirable in terms of payoff (cost) externality: if a player shares the cost of other players' breakdowns, he prefers them not to experiment in terms of direct costs. So the combination of positive informational externalities and negative payoff externalities is ambiguous *a priori*. Still, whether or not other players experiment too little or too much to a player's taste, he can nudge their

[12] The details are tedious, available from the authors.

action towards his preferred one by deviating accordingly when actions are observable. By playing safe, learning accelerates, prompting others to experiment; by playing risky, learning slows down, delaying experimentation by others. As a result, one would expect the equilibrium payoff and the amount of experimentation to be closer to their first-best levels when actions are observable, whether or not there is over-experimentation or not. This means, in particular, that whenever payoff externalities and informational externalities exactly offset each other in the unobservable case, if possible, so that first-best results, first-best should also obtain with observable actions.

We now confirm this discussion formally, by introducing payoff externalities in our model. This payoff externality has a simple interpretation as a cross-subsidization or insurance scheme: whenever a breakdown results, we assume that the agent who suffers this breakdown receives back a fraction $\alpha \leq 1$ as compensation, evenly shared by the other players (we assume that the identity of the player who suffers a breakdown is observable, whether actions are observable or not). Our baseline model corresponds to the special case $\alpha = 0$.

Formally, the total realized cost of player $i$ is now, given the realization of the process $\{N_t^i : t \geq 0\}_{i=1,\ldots,I}$,

$$\int_0^\infty re^{-rt}\left(h((\alpha/(I-1)\textstyle\sum_{j\neq i}\mathrm{d}N_t^j + (1-\alpha)\mathrm{d}N_t^i) + su_t^i\mathrm{d}t\right).$$

To stay away from equilibrium multiplicity at least under complete information, we assume throughout that

$$\alpha < \hat{\alpha} := \gamma\frac{\bar{u}+\mu-1}{\bar{u}+\mu}.$$

(This is a slightly stronger assumption than necessary, ruling out uninteresting complications.)

Let

$$\ell_\alpha^{**} := \ln\left[\frac{1+\mu}{\bar{u}+\mu-I}\frac{1}{\gamma-\alpha(1+\gamma)}\right],$$

and

$$\ell_\alpha^{*} := \ln\left[\frac{\bar{u}+\mu}{(\bar{u}+\mu)\left((1+\gamma)(1-\alpha)-1\right)-\gamma}\right].$$

We also introduce the following inequality that generalizes (11):

$$\alpha(1-\bar{u}+I)\frac{\gamma+1}{\gamma}(\bar{u}+\mu)+(\bar{u}-I)(\mu+\bar{u}) < \bar{u}-1. \tag{26}$$

We obtain the following theorem that generalizes Theorem 4 (Proofs are omitted and available from the authors).

**Theorem 10** *Consider the game with unobservable actions. There exists a unique symmetric equilibrium.*

1. *If* (26) *holds, this equilibrium involves mixed strategies. Specifically, player* $i$ *chooses a stopping strategy* $\pi_t$ *among the set* $[\underline{\tau}_\alpha, \bar{\tau}_\alpha]$, *with* $\underline{\tau}_\alpha < \bar{\tau}_\alpha$ *and* $\ell_{\bar{\tau}} = \ell^*_\alpha$; *this distribution is positive and continuous over* $(\underline{\tau}, \bar{\tau})$ *and has an atom at* $\bar{\tau}$. *The time* $\underline{\tau}$ *is strictly positive if and only if* $\ell^0 > \ell^{**}_\alpha$. *There is an atom at time* $\underline{\tau}_\alpha = 0$ *if and only if* $\ell^0 < \ell^{**}_\alpha$.
2. *If* (26) *does not hold, this equilibrium involves pure strategies. Specifically, player* $i$ *chooses* $\pi_{t^*}$, *where* $\ell_{t^*} = \ell^*_\alpha$.

It is easy to see the following.

**Lemma 5** *It holds that*

1. $\ell^{**}_\alpha > \ell^*_\alpha$ *if and only if* (26) *holds;*
2. *Condition* (26) *holds if and only if* $\alpha < \tilde{\alpha}$ *for some* $\tilde{\alpha} < \hat{\alpha}$.
3. $\ell^{**}_\alpha, \ell^*_\alpha$ *are strictly increasing and continuous in* $\alpha$ *(whenever* $\alpha < \hat{\alpha}$ *for* $\ell^*_\alpha$, *and whenever* $\alpha < \tilde{\alpha}$ *for* $\ell^{**}_\alpha$*).*
4. *There exists* $\alpha^*$ *such that* $\ell^*_{\alpha^*} = \ell^{FB}$. *Furthermore,* $\alpha^* \in (\tilde{\alpha}, \hat{\alpha})$.

The last item is particularly significant. *There exists a subsidy that makes players use the first-best cut-off as their unique equilibrium strategy.* The formula for the right subsidy is simple, namely

$$\alpha^* = \frac{\gamma(I-1)}{(1+\gamma)(I+\beta+\mu)}.$$

We now turn to the observable case. Define

$$\bar{\ell}_\alpha := \ell^*_\alpha - 1 - \frac{\bar{u} - I + 1}{\mu} - W_{-1}\left(-\tilde{\gamma}\frac{\bar{u} - I + \mu}{\mu} e^{\ell^*_\alpha - 1 - \frac{\bar{u} - I + 1}{\mu}}\right).$$

This threshold generalizes the threshold $\bar{\ell}$ introduced in Section 5.2 (that is, $\bar{\ell}_0 = \bar{\ell}$). It is easy to show that this threshold also is strictly increasing in $\alpha$.

We have the following generalization of Theorem 8 (we omit the specification of the precise amount assigned to the risky arm in the interior region).

**Theorem 11** *Consider the game with observable actions.*

1. *For* $\alpha \in [0, \alpha^*]$, *a unique symmetric Markov equilibrium exists. The safe arm is used for* $\ell \geq \bar{\ell}_\alpha$, *the risky arm for* $\ell \leq \ell^*_\alpha$, *and the amount assigned to the safe arm is continuously strictly increasing in the range* $\ell \in [\ell^*_\alpha, \bar{\ell}_\alpha]$.

2. *For $\alpha \in (\alpha^*, \hat{\alpha})$, multiple symmetric Markov equilibria exist, indexed by an (arbitrary) value $\ell_* \in [\ell_\alpha^B, \ell_\alpha^*]$ with $\ell_\alpha^B > \ell^{FB}$, such that the safe arm is used for $\ell \geq \ell_*$, and the risky arm for $\ell < \ell_*$.*

The multiplicity for $\alpha > \alpha^*$ reflects complementarities among the players' actions: if other players switch to the risky action, there are strict benefits to do as well. The choice $\ell_* = \ell_\alpha^B$ is the one for which smooth-pasting holds at the boundary.

The analysis bears out our discussion: for $\alpha > \alpha^*$, overexperimentation obtains whether actions are observed or not, but the extent of this over experimentation is worsened by the lack of observability. Similarly, for $\alpha < \alpha^*$, underexperimentation results, and lack of observability worsens it. Remarkably, for the right subsidy, whether actions are observed or not is irrelevant, and the unique symmetric (in the observable case, Markov) equilibrium achieves the first-best outcome.

Figure 8 illustrates these boundaries.

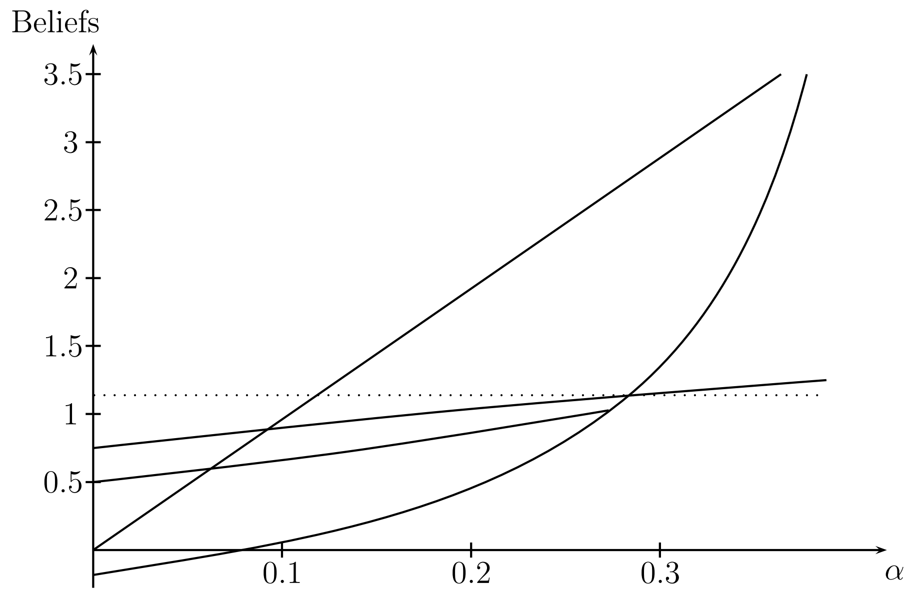



Figure 8: Pure- and Mixed-Strategy Equilibria with Subsidies, $(\bar{u}, I, \mu, \gamma) = (17/8, 2, 1/4, 2)$

# References


[1] Abreu, D., P. Milgrom and D. Pearce (1991). "Information and Timing in Repeated Partnerships," *Econometrica*, **59**, 1713–1733.

[2] Akcigit, U. and Q. Liu (2014). "The Role of Information in Innovation and Competition," working paper, Columbia University.

[3] Aumann, R.J. (1964). "Mixed and Behavior Strategies in Infinite Extensive Games," in *Advances in Game Theory*, Annals of Mathematics Studies Vol. 52, edited by M. Dresher, L.S. Shapley and A.W. Tucker, 627–650, Princeton University Press, Princeton.

[4] Board, S. and M. Meyer-ter-Vehn (2014). "A Reputational Theory of Firm Dynamics," working paper, UCLA.

[5] Bolton, P. and C. Harris (1999). "Strategic Experimentation," *Econometrica*, **67**, 349–374.

[6] Bonatti, A. and J. Hörner (2011). "Collaborating," *American Economic Review*, **101**, 632–663.

[7] Cesari, L. (1983). *Optimization: Theory and Applications. Problems with ordinary differential equations*, Applications of Mathematics, 17, Heidelberg: Springer-Verlag.

[8] Fleming, W.H. and H.N. Soner (2005). *Controlled Markov Processes and Viscosity Solutions*, Applications of Mathematics, 25, Heidelberg: Springer-Verlag.

[9] Holmström, B. (1999). "Managerial Incentive Problems: a Dynamic Perspective," *Review of Economic Studies*, **66**, 169–182. (Originally published in 1982 in Essays in Honor of Professor Lars Wahlbeck.)

[10] Karlin, S. (1959). *Mathematical Methods and Theory in Games, Programming and Economics*, Dover Publications.

[11] Keller, G. and S. Rady (2014). "Breakdowns," *Theoretical Economics*, forthcoming.

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

[13] Milgrom, P.R. and R.J. Weber (1985). "Distributional strategies for games with incomplete information," *Mathematics of Operations Research*, **10**, 619–632.

[14] Murto, P. and J. Välimäki (2011). "Learning and Information Aggregation in an Exit Game," *Review of Economic Studies*, **78**, 1426–1461.

[15] Presman, E.L. and I.N. Sonin (1990). *Sequential control with incomplete information*, in: Economic Theory, Econometrics, and Mathematical Economics, San Diego, CA: Academic Press.

[16] Rosenberg, D., A. Salomon and N. Vieille (2013). "On Games of Strategic Experimentation," *Games and Economic Behavior*, **82**, 31–51.

[17] Rothschild, M. (1974). "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, **9**, 185–202.

[18] Seierstad, A. and K. Sydsæter (1987). *Optimal Control Theory with Economic Applications*. North-Holland.

[19] Shmaya, E. and E. Solan (2014). "Equivalence between Random Stopping Times in Continuous Time," working paper, Kellogg School of Management.

[20] Thomas, C. (2014). "Strategic Experimentation with Congestion: Private Monitoring," working paper, University of Texas at Austin.

[21] Touzi, N. and N. Vieille (2002). "Continuous-time Dynkin Games with Mixed Strategies," *SIAM Journal of Control and Optimization*, **41**, 1073–1088.

[22] Vives, X. (1999). *Oligopoly Pricing, Old Ideas and New Tools*, Cambridge, MA: MIT Press.

[23] Weizsäcker, H. (1974). "Zur Gleichwertigkeit zweier Arten der Randomisierung," **11**, *Manuscripta Mathematica*, 91–94.

[24] Yushkevich, A.A. (1988). "On the two-armed bandit problem with continuous time parameter and discounted rewards," *Stochastics*, **23**, 299–310.

# Appendix

## A Proof of Lemma 1

See Keller and Rady 2014, Proposition 1 for the cooperative solution $u^{FB}$. Note that if $\ell > \ell^{FB}$, $u^{FB}(\ell) \leq u'(\ell) \leq u''(\ell)$ implies that $u^{FB}(\ell) = u'(\ell) = u''(\ell) = I$ and rewards are the same under all three policies until (if $\bar{u} > I$) $\ell < \ell^{FB}$. Hence, without loss, we assume $\ell^0 < \ell^{FB}$. Given some measurable $\underline{U}, \bar{U} : (-\infty, \ell^0] \to [0, I)$, with $0 \leq \underline{U}(\ell) < \bar{U}(\ell)$ and $\bar{U}$ bounded away from $I$, consider the program $\mathcal{P}^{FB}(\underline{u})$:

$$\min \int_t e^{-\mu t} \left( \mu \ell_t - \gamma(\bar{u} - 1 + \mu) e^{\ell_t} \right) \mathrm{d}t$$

over all $\pi : \mathbf{R}_+ \to [0, I]$, measurable, subject to

$$\dot{\ell}_t = u_t - \bar{u}, \quad \ell_0 = \ell^0.$$

with, for all $t \geq 0$ and $\ell_t \leq \ell^0$, $u_t \in [\underline{U}(\ell_t), \bar{U}(\ell_t)]$. By standard arguments, the optimal $u$ is measurable with respect to the belief $\ell$, and is the solution to the program

$$\min \int_\ell e^{-\mu t(\ell)} \left( \mu \ell - \gamma(\bar{u} - 1 + \mu) e^{\ell} \right) \mathrm{d}\ell,$$

over all $u : (-\infty, \ell^0] \to [0, I]$ such that $u(\ell) \in [\underline{U}(\ell), \bar{U}(\ell)]$, where $t(\ell)$ solves $t(\ell^0) = 0$ and

$$t'(\ell) = (u(\ell) - \bar{u})^{-1},$$

which is well-defined because $u(\ell) < \bar{U}(\ell) < I$. A routine application of the maximum principle (Theorem 4.2, Cesari 1983) yields that the optimal policy solves $u(\ell) = \underline{U}(\ell)$ a.e. Given $u', u''$ as stated in the lemma, the result follows if $u'' < I$ by setting $\underline{U} = u'$, $\bar{U} = u''$ and noting that $u''$ does not satisfy the necessary conditions. The same argument applies with $\bar{U} = I$ except in the case $\bar{u} = I$ where trivial modifications are needed in case $\lim_{t \to \infty} \ell''_t > -\infty$.

## B Reformulation of the Objective

Here we reformulate each player's objective, and we keep track of additional cost terms that will be necessary for comparative statics. Ignoring the uninsurable risk component $K$, each player minimizes

$$\int_{t \geq 0} e^{-rt} \left( r p_t g \left(1 - u_t^i\right) + r u_t^i s + \lambda p_t (\bar{u} - u_t^i - \nu_t^{-i}) s \right) \frac{1 - p^0}{1 - p_t} \mathrm{d}t, \tag{27}$$

subject to

$$\dot{p} = -\lambda p_t(1-p_t)(\bar{u} - u_t^i - \nu_t^{-i}).$$

Let us do the transformations slowly, first rewriting the objective in terms of the log-likelihood ratio $\ell_t := \ln(p_t/(1-p_t))$.

$$\int_{t\geq 0} e^{-rt}\left(re^{\ell_t}g(1-u_t^i) + ru_t^i s(1+e^{\ell_t}) + \lambda e^{\ell_t}(\bar{u} - u_t^i - \nu_t^{-i})s\right)(1+e^{\ell^0})^{-1}\mathrm{d}t.$$

Next, we make the change of variable $t \mapsto t/\lambda$, and we define $\gamma := (g-s)/s$ and $\mu := r/\lambda$. Finally, we factor out $(1+e^{\ell^0})^{-1}$.

$$\int_{t\geq 0} e^{-\mu t}\left(\mu e^{\ell_t}g + \mu(s(1+e^{\ell_t}) - ge^{\ell_t})(\dot{\ell}_t + \bar{u} - \nu_t^{-i}) - \dot{\ell}_t e^{\ell_t}s\right)\mathrm{d}t.$$

Integrating the last term yields

$$e^{\ell^0}s + \int_{t\geq 0} e^{-\mu t}\left(e^{\ell_t}(\mu g + \mu(s-g)(\dot{\ell}_t + \bar{u} - \nu_t^{-i})) + \mu s(\dot{\ell}_t + \bar{u} - \nu_t^{-i}) - \mu s e^{\ell_t}\right)\mathrm{d}t.$$

Integrating the first two terms by parts, and factoring out $s$, we obtain the following expression to the expected cost

$$W(\ell^0) := \frac{s(1+\mu\gamma)}{1+e^{-\ell^0}} + \frac{\mu s}{1+e^{\ell^0}}\int_{t\geq 0} e^{-\mu t}\left(\mu(\ell_t - \ell^0) - \gamma(\bar{u} - \nu_t^{-i} - 1 + \mu)e^{\ell_t} + \bar{u} - \nu_t^{-i}\right)\mathrm{d}t. \tag{28}$$

Therefore, eliminating constant terms, player $i$ minimizes

$$\int_{t\geq 0} e^{-\mu t}\left(\mu\ell_t - \gamma(\bar{u} - \nu_t^{-i} - 1 + \mu)e^{\ell_t}\right)\mathrm{d}t,$$

subject to

$$\dot{\ell}_t = u_t^i + \nu_t^{-i} - \bar{u}.$$

# C Proof of Lemma 2

The proof of this lemma relies on the proof of Lemma 3, proved next and independently (except for the last sentence of that next proof, which isn't used here).

We apply the maximum principle to $\mathcal{P}$. The maximum principle implies that there exists an absolutely continuous $\psi : \mathbf{R}_+ \to \mathbf{R}$ such that (i) $\psi_t > 0 \Rightarrow u_t^i = 0$, (ii) $\psi_t < 0 \Rightarrow u_t^i = 1$, and (iii) almost everywhere

$$\dot{\psi}_t e^{\mu t} = \gamma(\bar{u} - \nu_t^{-i} - 1 + \mu)e^{\ell_t} - \mu.$$

Because $\nu_t^{-i} \leq I-1$, a sufficient condition for $\dot{\psi}_t > 0$ at any time $t$ such that $\ell_t \geq \ell^*$ is that

$$\gamma(\bar{u}-(I-1)-1+\mu)e^{\ell^*} > \mu.$$

Using the definition of $\ell^*$, this is equivalent to

$$(\mu+\bar{u})(\mu+\bar{u}-I) \geq \mu(\mu+\bar{u}-1),$$

or

$$\bar{u}(\bar{u}-I)+\mu(\bar{u}-I+1) \geq 0,$$

which is true.

It follows that $\psi$ is strictly increasing at all times $t$ such that $\ell_t \in [\ell^*, \ell^0]$; hence, given (i), there exists $\bar{t} \geq 0$ such that any solution must specify $u_t^i = 1$ for all $t < \bar{t}$ and $u_t^i = 0$ for $t \geq \bar{t}$ (recall that $u_t^i = 0$ when $\ell_t < \ell^*$).

# D Proofs for Section 3.4

Suppose player $j \neq i$ switches to the risky arm at time $T$. Then $i$ chooses $\tau$ to minimize

$$\begin{aligned} &\int_{t\leq T} e^{-rt}\left(\mu s(\ell+\lambda(t\wedge\tau+(I-1-\bar{u})t)) - (g-s)(\bar{u}-I+\mu)e^{\ell+\lambda(t\wedge\tau+(I-1-\bar{u})t)}\right)\mathrm{d}t \quad (29)\\ +&\int_{t\geq T} e^{-rt}\left(\mu s(\ell+\lambda(t\wedge\tau+(I-1)T-\bar{u}t)) - (g-s)(\bar{u}-1+\mu)e^{\ell+\lambda(t\wedge\tau+(I-1)T-\bar{u}t)}\right)\mathrm{d}t.\end{aligned}$$

Integrating (29) gives, for $\tau \leq T$,

$$\begin{aligned}\mathcal{C}^L(x^i,x^{-i}) = \; & ze^{-x^i(\mu+\bar{u}-I)+\ell} - \frac{z(\mu+\bar{u}-1)e^{-x^{-i}(\mu+\bar{u}-I+1)+x^i+I}}{\mu+\bar{u}} - \frac{z(I-\mu-\bar{u})e^{-x^{-i}(\mu+\bar{u}-I+1)+x^i+\ell}}{\mu+\bar{u}-I+1}\\ &+\frac{z(I-\mu-\bar{u})e^{\ell-x^i(\mu+\bar{u}-I)}}{\mu+\bar{u}-I+1} - \frac{(I-1)e^{-\mu x^{-i}}}{\mu} - \frac{e^{-\mu x^i}}{\mu} + \frac{I-\bar{u}}{\mu} - e^{\ell}z+\ell,\end{aligned}$$

where $x^i = \tau/\lambda$, $x^{-i} = T/\lambda$. For $\tau \geq T$, we get

$$\mathcal{C}^F(x^i,x^{-i}) = \frac{e^{-x^i(\mu+\bar{u})-(\mu+1)x^{-i}}(K_1+K_2)}{\mu(\mu+\bar{u})},$$

where

$$K_1 := (\mu+\bar{u})e^{\bar{u}x^i+x^{-i}}\left(e^{\mu\left(x^{-i}+x_i\right)}(I+\mu\ell-\bar{u}) - (I-1)e^{\mu x^i} - e^{\mu x^{-i}}\right)$$

and

$$K_2 := \mu z e^{\mu x_{-i}+\ell}\left(e^{Ix_{-i}+x_i} - (\mu+\bar{u})e^{x^i(\mu+\bar{u})+x^{-i}}\right).$$

We recall the claims from Section 3.4, expressed here equivalently in terms of $x^i, x^j$.

**Claim 12** *The minimizer $x^i \geq x^{-i}$ of $\mathcal{C}^S$ is increasing in $x^{-i}$.*

**Proof.** By direct differentiation, the sign of $\mathrm{d}\mathcal{C}^S/\mathrm{d}x^i$ is the same as the sign of

$$1 - \frac{\mu + \bar{u} - 1}{\mu + \bar{u}} z e^{\ell - (\bar{u}-1)x^i + (I-1)x^{-i}},$$

so that if the minimum is interior, $x^i$ increases when $x^{-i}$ does. [The second derivative is of the sign of

$$\frac{(\mu + \bar{u} - 1)^2 z}{\mu + \bar{u}} e^{\ell - (\bar{u}-1)x + (I-1)x^{-i}} - \mu,$$

hence it is (at most) first positive (so that $V^S$ is convex) and then negative (so that $\mathcal{C}^F$ is concave). But note that the sign of $\mathrm{d}\mathcal{C}^S/\mathrm{d}x^i$ is positive as $x^i \to \infty$, establishing that there is at most one critical point, which must be a minimum.] Note that this critical point is strictly greater than $x^{-i}$ if and only if $x^{-i}$ is less than

$$\hat{x}^{-i} := \frac{\ell + \ln \frac{(\mu+\bar{u}-1)z}{\mu+\bar{u}}}{\bar{u} - I},$$

above which the minimum is achieved by $x^i = x^{-i}$, which is clearly increasing as well. ■

**Claim 13** *If the minimizer of $\mathcal{C}^F$ is $x^i < x^{-i}$, then it is decreasing in $x^{-i}$ at $x^i$.*

**Proof.** By direct differentiation, the sign of $\mathrm{d}\mathcal{C}^F/\mathrm{d}x^i$ is the same as

$$f(x^i, x^{-i}) := (\mu+\bar{u})(1-I+\mu+\bar{u}) - z(I-1)e^{\ell+(1+\mu+\bar{u})(x^i-x^{-i}))+Ix^{-i}-\bar{u}x^i} - (\mu+\bar{u})(\mu+\bar{u}-I)ze^{\ell+(I-\bar{u})x^i},$$

which (again by direct differentiation) is concave. This implies that $\mathcal{C}^F$ is (at most) first decreasing, then increasing, then decreasing. This implies that there is at most one local interior minimum, and that either this local minimum is the global minimum, or the global minimum is $x^i = x^{-i}$ (note that this cannot be the case if $\mathrm{d}\mathcal{C}^F/\mathrm{d}x^i > 0$ at $x^i = x^{-i}$). Consider now the local minimum. Let $g(x^i, x^{-i}) = e^{-(1+\mu)x^i} f(x^i, x^{-i})$, and note that the local minimum is a zero of $g$. Computing,

$$-\frac{\frac{\partial g(x^i,x^{-i})}{\partial x^i}}{\frac{\partial g(x^i,x^{-i})}{\partial x^{-i}}} = \frac{(\mu + \bar{u})}{(I-1)z}\left((1+\mu) - (\mu + \bar{u} - I)ze^{\ell-(\bar{u}-I)x^i}\right)e^{-\ell-(1+\mu)x^i+(1+\mu+\bar{u}-I)x^{-i}}.$$

We now note that

$$\left.\frac{\mathrm{d}\mathcal{C}^F(x^i, x^{-i})}{\mathrm{d}x^i}\right|_{x^i=x^{-i}} = \frac{e^{-\mu x^{-i}}}{\mu + \bar{u}}\left((\mu + \bar{u}) - (\mu + \bar{u} - I)ze^{\ell-(\bar{u}-I)x^{-i}}\right),$$

which is zero at $\hat{x}^{-i}$. Hence (as $\bar{u} \geq 1$), the interior minimum is decreasing as long as $x^i \leq \hat{x}^{-i}$. (Note that this implies that this minimum is either equal to $\hat{x}^{-i}$ at $x^{-i}$, or strictly lower, depending upon whether $\mathcal{C}^F$ is concave there; in any event, it is well-defined at any point sufficiently close but to the right of $\hat{x}^{-i}$.) But since $x^i$ is below the diagonal (that is, $x^i < x^{-i}$ by assumption), and so below $\hat{x}^{-i}$ for $x^{-i} = \hat{x}^{-i}$ (formally, strictly for all strictly higher $x^{-i}$), it is lower than $\hat{x}^{-i}$ for all $x^{-i}$. Hence it is decreasing for all $x^{-i}$ at which is well-defined. ■

**Claim 14** *There exists* $\bar{x}^{-i}$ *such that* $\min_{x^i \geq x^{-i}} \mathcal{C}^S(x^i, x^{-i}) < \min_{x^i < x^{-i}} \mathcal{C}^F(x^i, x^{-i})$ *if and only if* $x^{-i} < \bar{x}^{-i}$.

**Proof.** Note that this is immediate when (11) is violated (that is, when a pure-strategy equilibrium exists). This is because, for $x^i > \bar{x}^{-i}$, the only candidate for a global minimum over $\{x^i : x^i \geq x^{-i}\}$ is $x^i = x^{-i}$. Yet $\mathcal{C}^F$ is increasing at $x^i = x^{-i}$ for $x^{-i} > \bar{x}^{-i}$, hence the global minimum for $x^{-i} > \bar{x}^{-i}$ must be the local minimum of $\mathcal{C}^F$, which is then well-defined. As for $x^i < \bar{x}^{-i}$, the local minimum of $\mathcal{C}^L$ is no longer well-defined (as it hits the diagonal at $\bar{x}^{-i}$); so the unique candidate at $x^i < \bar{x}^{-i}$ is the local minimum $x^i > \bar{x}^{-i}$ of $\mathcal{C}^S$, and indeed, $\mathcal{C}^S$ is then decreasing at $x^i = x^{-i}$, so that this local minimum is well-defined.

So we may as well assume that (11) is satisfied, and it is clear that the global minimum must be the local minimum of $\mathcal{C}^F$ for $x^i > \bar{x}^{-i}$. Similarly, there is nothing to show if $V^F$ no local minimum $x^i < x^{-i}$ (which, by the previous claim, is equivalent to $x^{-i}$ exceeding some threshold). Now, by the envelope theorem, the difference in costs between the local minima of $\mathcal{C}^S$ and $\mathcal{C}^F$ is given by

$$\frac{\partial \mathcal{C}^F(x^i, x^{-i})}{\partial x^{-i}} - \frac{\partial V_2(x^i, x^{-i})}{\partial x^{-i}} = \frac{(I-1)z}{\mu + \bar{u}} e^{\ell - (1+\mu+\bar{u}-I)x^{-i}} \left( e^{x^F} - e^{x^S - (\mu+\bar{u})(x^S - x^{-i})} \right),$$

where $x^k$ is the local minimum of $\mathcal{C}^k$, $k = F, S$. So the result follows if we can show that

$$x^{-i} - x^F - (\mu + \bar{u} - 1)(x^S - x^{-i}) \geq 0,$$

Given that solving $\frac{\mathrm{d}\mathcal{C}^S}{\mathrm{d}x^i} = 0$ gives

$$x^S = \frac{\ell + (I-1)x^{-i} + \ln \frac{(\mu+\bar{u}-1)z}{\mu+\bar{u}}}{\bar{u} - 1},$$

we have that $x^S - x^{-i}$ is decreasing (in $x^{-i}$), while $x^{-i} - x^F$ is increasing. So $\mathcal{C}^F$ and $\mathcal{C}^S$ cross at most once, but when $x^F$ hits the diagonal $x^i = x^{-i}$, we must have $\mathcal{C}^S < \mathcal{C}^F$, and so the result follows. ■

# E Proof of Proposition 3

The second inequality of the proposition ($\nu^o(p) \geq \nu^{fb}(p)$) being immediate given that $\bar{p} < p^{FB}$, it is the first inequality that must be established. Given $\ell^0$ and $\ell < \ell^0$, we let $t(\ell)$ denote the time at which the belief of the outside observer reaches belief $\ell$. We normalize $\ell^0 = \ell^{**}$ so that $t(\ell^{**}) = 0$. Further, let $\nu_t^i$ denote the hazard rate of the outsider's belief at time $t$, *i.e.*, his belief satisfies

$$\dot{\ell}_t = I\nu_t^i - \bar{u}, \quad \ell_0 = \ell^0.$$

Now suppose towards a contradiction that there exists a belief level $\hat{\ell}$ such that the outside observer's hazard rate in the unobservable case $\nu^n(\hat{\ell})$ is equal to the hazard rate in the observable case $\nu^o(\hat{\ell})$. We derive an ordinary differential equation for $\nu^{-i}(\ell) := (I-1)\nu_{t(\ell)}^i$ in both cases.

In the unobservable case, we know from Theorem 4 that

$$\nu_t^{-i} = -1 + \bar{u} + \frac{\mu}{1 - \frac{e^{-\mu t}(-I+\mu+\bar{u})}{\bar{u}-I}}.$$

Differentiating $\nu_t$ with respect to $t$, we obtain

$$\frac{\mathrm{d}\nu_t^{-i}}{\mathrm{d}t} = -\frac{e^{\mu t}\mu^2(\bar{u} - I)(\bar{u} + \mu - I)}{(-I + \mu - e^{\mu t}(\bar{u} - I) + \bar{u})^2}.$$

Solving for $e^{\mu t}$ from the definition of $\nu_t^{-i}$ and plugging back into the derivative, we obtain

$$\frac{\mathrm{d}\nu^{-i}(\ell)}{\mathrm{d}\ell} = -(-1 + \bar{u} - \nu(\ell))(-1 + \mu + \bar{u} - \nu(\ell))t'(\ell),$$

where

$$t'(\ell) = \frac{1}{\frac{I}{I-1}\nu^{-i}(\ell) - \bar{u}}.$$

Finally, we obtain the equation,

$$\frac{\mathrm{d}\nu^{-i}(\ell)}{\mathrm{d}\ell} = (\mu + \bar{u} - \nu^{-i}(\ell) - 1)\frac{\bar{u} - \nu^{-i}(\ell) - 1}{\bar{u} - \nu^{-i}(\ell) - \frac{\nu^{-i}(\ell)}{I-1}}. \tag{30}$$

Note that $\nu^{-i}(\ell)$ is increasing in $\ell$, as expected. Also notice that the second term in (30) is smaller than one, because $\nu^{-i} \leq I - 1$.

In the observable case, we already have the expression for the hazard rate

$$\nu^{-i}(\ell) = \mu + \bar{u} - 1 - \frac{1 + (\ell - \ell^*)\mu}{e^\ell\gamma - 1}.$$

Differentiating with respect to $\ell$ and replacing $e^{\ell}$ with the solution to the previous equation, we obtain the following differential equation

$$\frac{\mathrm{d}\nu^{-i}(\ell)}{\mathrm{d}\ell} = (\mu + \bar{u} - \nu^{-i}(\ell) - 1)\frac{\bar{u} - \nu^{-i}(\ell) + \mu(\ell - \ell^{*})}{1 + \mu(\ell - \ell^{*})} \tag{31}$$

Notice that $\bar{u} - \nu^{-i} > \bar{u} - I + 1 > 1$, and therefore in the ratio in (31) is larger than one. Furthermore, the first term $(\mu + \bar{u} - \nu^{-i} - 1)$ is identical in the two expressions (30) and (31). Thus, if the two paths $\nu^{o}(\ell)$ and $\nu^{n}(\ell)$ cross, the observable path must be steeper. This yields a contradiction, because

$$\nu^{o}(\ell^{**}) < \nu^{n}(\ell^{**}) = I - 1,$$

and therefore if the paths $\nu(\ell)$ cross, the unobservable path must be steeper at the crossing point closest to $\ell^{**}$.