Hopeful Experimentation

Faheem H. Gilani*

Working draft: October 10, 2026

Abstract

Is a decision maker more or less persistent when she does not know whether a project, if viable, pays large lump sums rarely or small ones often? I study an exponential bandit in which a viable project has an unknown arrival rate but a fixed expected flow payoff. A decision maker should continue while her belief that the project is viable exceeds the known-rate cutoff at the current mean rate conditional on viability. Relative to a known-rate benchmark with the same mean, a lack of breakthroughs erodes this belief more slowly, since it suggests that a viable project pays rarely but handsomely (*hope*), but raises the cutoff by lowering the conditional mean rate (*discouragement*). Discouragement dominates early and hope later: absent a breakthrough, this rate–size uncertainty makes her give up sooner on less promising projects and later on more promising ones. Whenever experimentation is worthwhile, it also makes her worse off, more likely to abandon a viable project, and more optimistic about viability when she gives up. With several players, the unique symmetric Markov perfect equilibrium has them experiment fully while sufficiently optimistic and then mix at an explicit, declining intensity that keeps each indifferent between experimenting and free riding; beliefs converge to the single-agent boundary without reaching it.

Keywords: Experimentation; Bayesian learning; Bandits; Optimal stopping; Free riding.
JEL classification: C73; D81; D83.

*PhD student, Department of Economics, The University of Texas at Austin, Email: fhg3@utexas.edu.

1 Introduction

A novelist who has published nothing in five years may lack talent, or she may write one great book a decade. A patron deciding whether to keep supporting her faces two uncertainties: whether she can succeed and, if so, whether she produces rare great books or frequent modest ones. In the now canonical exponential bandit model of Keller et al. (2005), a good risky arm produces breakthroughs at a known Poisson rate, so a lack of success (*silence*) is unambiguously bad news about viability. This paper introduces uncertainty about the arrival rate of a good risky arm while keeping the expected flow payoff the same. Thus a slow good arm pays large lump sums rarely while a fast one pays small sums often. This *rate–size uncertainty* separates learning about when and how much a viable project pays from learning about its long-run profitability.

A first payoff resolves the decision maker's (DM's) uncertainty about viability, but silence has two effects on the posterior. It lowers the probability that the arm is good, but among good arms it shifts probability toward slow ones. Each further moment of silence is therefore less damning than the last, a force I call *hope*.[1] When the prior puts enough weight on fast arms, the probability that the arm is "good but slow" can even rise without news. The counterpart is *discouragement*: the declining mean arrival rate conditional on viability reduces the prospect of an early breakthrough.

I first analyze the single-player decision problem and then characterize the symmetric Markov perfect equilibrium in the N-player game. In this model the state variable (posterior belief about viability and the posterior distribution of arrival rates) is multidimensional, but after a history without news it depends only on cumulative experimentation, the *experimentation clock*. Payoffs depend on the distribution of arrival rates through its Laplace transform. This reduction yields a tractable single-agent problem and turns strategic experimentation into a game of dynamic contributions to a public good (Lemmas 2 and 3).

I show that for any distribution of arrival rates with finite mean, the DM continues experimenting while her belief in viability exceeds the known-rate cutoff evaluated at the *current* mean arrival rate conditional on viability (Theorem 1). This bar rises as silence accumulates. The optimal rule is myopic because the net benefit of experimentation declines along every no-news path. The analysis also gives a linear stopping boundary for finitely many arrival rates, a scalar characterization of the stopping clock and value, and an explicit Gittins index for allocating effort among independent projects (Corollary 3).

[1] The term is unrelated to anticipatory utility (e.g., Caplin and Leahy, 2001).

I then compare this model to an exponential bandit benchmark where the rate is the prior mean arrival rate conditional on viability, and where the prior probability of viability is the same. I show that uncertainty leaves the decision to start unchanged and, whenever experimentation is worthwhile, it lowers value, leads to abandonment at more optimistic beliefs about viability, and raises the probability of abandoning a good arm (Proposition 1). Value is also nonincreasing in dispersion in the convex order, with a strict decline when experimentation is worthwhile. The effect on persistence depends on the project: uncertainty shortens the *abandonment horizon*, the time the DM is willing to experiment without a breakthrough, for less promising projects and lengthens it for more promising ones (Theorem 2). Discouragement dominates early, as fast types are ruled out, and hope later, as slow types prove hard to rule out. Because uncertainty also changes breakthrough timing, expected time until success or abandonment need not move in the same direction (Remark 2).

The N-player game is set up similarly to the one in Keller et al. (2005), in which players have replica arms with common viability and a common arrival rate, and where the history of experimentation and payoffs is public. I characterize efficient experimentation and the unique symmetric Markov perfect equilibrium in admissible strategies (Proposition 3 and Theorem 3). Sufficiently optimistic players experiment fully; they then mix at a declining intensity that keeps each indifferent between experimenting and free riding. If experimentation starts, beliefs approach the single-agent boundary asymptotically, and effort is inefficiently low near that boundary. The single-agent comparisons for terminal beliefs and the probability of abandoning a good project therefore carry over to teams. Uncertainty also reduces the *proportional* loss of good projects caused by free riding, although numerical examples show that its effect on the welfare cost of free riding can have either sign; that cost vanishes in the extreme-uncertainty limit of Proposition 4(c) (see also Fig. 5).

1.1 Related literature

The paper contributes to the literature on bandit models and strategic experimentation (Bolton and Harris, 1999; Keller et al., 2005); see Bergemann and Välimäki (2008) and Hörner and Skrzypacz (2017) for surveys. The closest benchmark is the conclusive-good-news model of Keller et al. (2005), obtained here when the distribution of arrival rates is degenerate. In Keller and Rady (2010), the risky arm pays lump sums with a common mean at one of two positive rates, so rate and quality coincide. Here viability and arrival rate are separate dimensions, and the latter is payoff-irrelevant once viability is known.

Cohen and Solan (2013) allow the risky arm to follow one of two Lévy processes; here it follows a compound Poisson process with an arbitrary prior over intensity. Multidimensional beliefs also arise when a DM learns about several hypotheses from Poisson signals (Mayskaya, 2024); the clock reduction here exploits the one-dimensional no-news paths of a single risky arm.

Learning about hazard rates appears in R&D races (Malueg and Tsutsui, 1997; Choi, 1991), job search (Gonzalez and Shi, 2010), and tasks with uncertain work requirements (Ely and Szydlowski, 2020). Boyarchenko (2020) studies stopping games with non-exponential news arrival and the effects of changing hazards on strategic interaction. Külpmann (2026) studies teams with an unknown effort requirement and a deadline, where unobservable effort and quadratic costs link learning, free riding, and the timing of work. Here players observe one another's experimentation, and learning concerns viability and a persistent common arrival rate. The fixed expected flow payoff of a good arm isolates the effects of rate–size uncertainty.

Mixtures of exponential survival functions generate declining hazards through dynamic selection, a central theme of duration analysis (Heckman and Singer, 1984; Lancaster, 1990). Related mathematics underlies hyperbolic discounting from uncertain hazards (Sozou, 1998) or realization times (Dasgupta and Maskin, 2005), and discounting the far future at the lowest possible rate (Weitzman, 1998). Here the lowest arrival rates govern long horizons and allow hope to dominate. The rising abandonment bar also parallels the collapsing decision boundaries in Fudenberg et al. (2018) when decision stakes are unknown.

The strategic analysis uses the analogy with dynamic contributions to a public good (Admati and Perry, 1991; Marx and Matthews, 2000; Georgiadis, 2015; Campbell et al., 2014). The single-agent result is an instance of the "monotone case" of optimal stopping (Chow et al., 1971), and the index is a Gittins index (Gittins, 1979; Gittins et al., 2011; El Karoui and Karatzas, 1994).

2 Model

2.1 Environment

Time $t \in [0, \infty)$ is continuous and payoffs are discounted at rate $r > 0$. A decision maker (DM) allocates one unit of a perfectly divisible resource per unit of time between a safe arm S and a risky arm R. Let $k_t \in [0, 1]$ be the fraction allocated to R at time t. The safe

arm yields the deterministic flow payoff $(1-k_t)s$. The risky arm is characterized by an unknown *arrival rate* $\Lambda \geq 0$ per unit of experimentation. Given Λ, lump-sum payoffs arrive according to a Poisson process with intensity $k_t\Lambda$, and their sizes are independent draws, independent of the arrival times, with mean $h(\Lambda)$. The arm is *bad* if $\Lambda = 0$, in which case it never pays, and *good* (or viable) with rate λ if $\Lambda = \lambda > 0$.

Assumption 1 (Rate–size trade-off). (i) $h(\lambda) = g/\lambda$ for all $\lambda > 0$, where $g > s > 0$.

(ii) $\Pr(\Lambda > 0) = p_0 \in (0,1)$ and, conditional on $\Lambda > 0$, Λ is distributed according to a Borel probability measure G on $(0,\infty)$ with finite mean $\hat{\lambda} := \int \lambda \, G(\mathrm{d}\lambda)$.

Under Assumption 1(i), every good arm yields the expected flow payoff $g > s$ per unit of resource, so a DM who knew the arm to be good would use it exclusively. The DM is risk neutral and maximizes

$$\mathbb{E}\left[\int_0^\infty re^{-rt}\big((1-k_t)s\,\mathrm{d}t + \mathrm{d}\Pi_t\big)\right],$$

where Π_t is the cumulative sum of lump sums received up to time t. The factor r expresses values in flow units. Because k_t may depend only on information available before t, the expected payoff equals $\mathbb{E}\big[\int_0^\infty re^{-rt}\big((1-k_t)s + k_t g\,\mathbf{1}\{\Lambda > 0\}\big)\mathrm{d}t\big]$.

The DM observes the arrival times and sizes of lump sums, or *breakthroughs*. A first breakthrough reveals that the arm is good, after which the DM optimally uses R forever and earns g. Before the first breakthrough she observes nothing but the passage of time, so a policy is fully described by a measurable function $k : [0,\infty) \to [0,1]$ specifying the allocation in the absence of a breakthrough. Randomizing over policies cannot help, because the payoff of a randomized policy is an average of the payoffs of deterministic ones.

Two special cases organize the discussion. If $G = \delta_\lambda$ is degenerate, the model is the exponential bandit of Keller et al. (2005), whose single-agent cutoff will play a central role:

$$\alpha(\lambda) := \frac{rs}{rg + \lambda(g-s)}. \tag{1}$$

In this *known-rate benchmark*, the arrival rate λ is known conditional on viability, but viability remains uncertain. The DM experiments if and only if her belief that the arm is good exceeds $\alpha(\lambda)$. If G puts mass on two arrival rates $0 < \lambda^- < \lambda^+$, it is convenient to describe beliefs by the pair (p^-, p^+) of probabilities that the arm is good and slow, and good and fast, a point in the triangle $\{p^-, p^+ \geq 0,\ p^- + p^+ \leq 1\}$.

2.2 Learning from silence

Let $T_t := \int_0^t k_u \, \mathrm{d}u$ be the cumulative amount of experimentation by time t, the *experimentation clock*. Conditional on Λ, the probability of no arrival by time t is $e^{-\Lambda T_t}$. The posterior after a history without news therefore depends on the history only through T_t, and every no-news belief can be indexed by the clock. Let

$$L(T) := \int_{(0,\infty)} e^{-\lambda T} \, G(\mathrm{d}\lambda), \qquad D(T) := 1 - p_0 + p_0 L(T). \tag{2}$$

Here $L(T)$, the Laplace transform of G, is the probability that a good arm stays silent for T units of experimentation, and $D(T)$ is the unconditional probability of silence. The function L is smooth on $(0,\infty)$, strictly decreasing, and strictly convex, with $L(0) = 1$ and $L'(0) = -\hat{\lambda}$. Denote by p_T the posterior probability that the arm is good, by G_T the posterior distribution of its arrival rate conditional on being good, by m_T and v_T the mean and variance of G_T, and by

$$\bar{\lambda}_T := p_T m_T = \mathbb{E}[\Lambda \mid \text{silence for } T]$$

the posterior expected arrival rate, including the possibility that the arm is bad. For the known-rate benchmark with rate $\hat{\lambda}$, write $p_T^\delta := p_0 e^{-\hat{\lambda} T} / (1 - p_0 + p_0 e^{-\hat{\lambda} T})$.

Lemma 1 (Learning from silence). *(a) $p_T = p_0 L(T)/D(T)$, and the posterior distribution of a good arm's arrival rate and its mean are*

$$G_T(\mathrm{d}\lambda) = \frac{e^{-\lambda T} \, G(\mathrm{d}\lambda)}{L(T)}, \qquad m_T = -\frac{L'(T)}{L(T)}.$$

(b) For $T > 0$ (and at $T = 0$ if G has finite variance), the odds that the arm is good decay at rate m_T, and the derivative of the mean arrival rate conditional on viability equals minus its variance:

$$\frac{\mathrm{d}}{\mathrm{d}T} \ln \frac{p_T}{1 - p_T} = -m_T, \qquad \frac{\mathrm{d}m_T}{\mathrm{d}T} = -v_T.$$

(c) (Replicator dynamics.) *For every Borel set $A \subseteq [0,\infty)$ of arrival rates, $\frac{\mathrm{d}}{\mathrm{d}T} \Pr(\Lambda \in A \mid T) = \mathbb{E}\big[(\bar{\lambda}_T - \Lambda)\mathbf{1}\{\Lambda \in A\} \mid T\big]$. In particular, if Λ takes finitely many values, the posterior probability of each value rises if and only if it is below the expected arrival rate $\bar{\lambda}_T$. Moreover, $\mathrm{d}\bar{\lambda}_T/\mathrm{d}T = -\operatorname{Var}(\Lambda \mid T)$ for $T > 0$.*

(d) (Hope and discouragement.) *If G is not degenerate, then $p_T > p_T^\delta$ and $m_T < \hat{\lambda}$ for all $T > 0$.*

All proofs are in Appendix A. Part (b) captures hope and discouragement: the declining conditional mean rate makes silence progressively less informative about viability while reducing the prospect of a quick breakthrough. Part (c) is a replicator dynamic with fitness equal to minus the arrival rate; its variance identity is the corresponding version of Fisher's fundamental theorem (see Hofbauer and Sigmund, 1998).

Corollary 1 (Two arrival rates; the tipping line)**.** *Let G put mass on $\lambda^- < \lambda^+$. Along a no-news path, p_T^+ and $p_T = p_T^- + p_T^+$ strictly decrease, while p_T^- strictly increases when $\lambda^- p_T^- + \lambda^+ p_T^+ > \lambda^-$ and strictly decreases when $\lambda^- p_T^- + \lambda^+ p_T^+ < \lambda^-$. Since $\bar{\lambda}_T$ declines, p_T^- is single-peaked along every no-news path: it rises while the path lies above the* tipping line *$\lambda^- p^- + \lambda^+ p^+ = \lambda^-$ and falls once the path is below it (from the outset if the prior lies on or below the line).*

Learning about the arrival rate may continue after a breakthrough, as in Keller and Rady (2010), but, since $g > s$ by Assumption 1, this learning does not affect behavior.

2.3 Payoffs as a function of the clock

Two functions of the clock summarize the economics of experimenting. Let

$$\chi(T) := (1 - p_0)s - (g - s)p_0 L(T) = D(T)\left(s - g p_T\right), \tag{3}$$

$$\psi(T) := \chi'(T) - r\chi(T) = (g - s)p_0 \int (r + \lambda) e^{-\lambda T} G(\mathrm{d}\lambda) - r(1 - p_0)s. \tag{4}$$

Seen from time zero, $\chi(T)$ is the expected flow loss from playing R instead of S at clock T before a breakthrough: the probability $D(T)$ of reaching clock T in silence, times the myopic loss $s - g p_T$. Its derivative, $\chi'(T) = (g - s)p_0 \int \lambda e^{-\lambda T} G(\mathrm{d}\lambda)$, is the gain $g - s$ times the ex ante probability density that the first breakthrough occurs at clock T. Factoring out $D(T)$ gives the normalized form

$$\psi(T) = D(T)\left[(g - s)\,\bar{\lambda}_T - r\left(s - g p_T\right)\right]. \tag{5}$$

The bracket compares the expected capital gain from a breakthrough, which arrives at rate $\bar{\lambda}_T$ and raises the continuation value from s to g, with the flow cost of experimenting multiplied by the discount rate.

Lemma 2 (Payoff representation)**.** *For every policy k, with clock $T_t = \int_0^t k_u \, \mathrm{d}u$, the DM's expected payoff is*

$$J(k) = s + \int_0^\infty e^{-rt} \, \psi(T_t) \, k_t \, \mathrm{d}t \;=\; s + \int_0^\infty e^{-rt} \, \psi(T_t) \, \mathrm{d}T_t. \tag{6}$$

Every unit of experimentation performed at calendar time t, when the clock reads T_t, contributes $e^{-rt}\psi(T_t)$. The representation thus separates the two ways a policy matters: *which* clock values are visited, which determines $\psi(T_t)$, and *when*, which determines e^{-rt}. The distribution of arrival rates matters only through the Laplace transform L.

Remark 1 (Characteristics). In the two-rate model, the no-news belief paths are the characteristics of the Hamilton–Jacobi–Bellman equation on the belief triangle. Parameterizing them by the clock reduces that partial differential equation to a scalar ordinary differential equation, which Lemma 2 integrates.

3 Optimal experimentation

Theorem 1 (Optimal experimentation)**.** *Let Assumption 1 hold.*

(a) The function ψ is strictly decreasing, with $\lim_{T\to\infty} \psi(T) = -r(1-p_0)s < 0$. Let $T_1^* := \inf\{T \geq 0 : \psi(T) \leq 0\}$. Then $T_1^* < \infty$, and $T_1^* > 0$ if and only if $p_0 > \alpha(\hat{\lambda})$.

(b) It is optimal to experiment exclusively until a breakthrough occurs or the clock reaches T_1^*, whichever comes first, and to use the safe arm forever if no breakthrough has occurred by then. Every optimal policy coincides with this one at almost every time.

(c) Equivalently, at every belief the DM experiments if and only if

$$(g-s)\,\bar{\lambda} > r\,(s - gp) \qquad \Longleftrightarrow \qquad p > \alpha(m), \tag{7}$$

where p is the current probability that the arm is good, m is the current mean arrival rate conditional on viability, and $\bar{\lambda} = pm$.

(d) The value at the prior is

$$V_1 = s + \int_0^{T_1^*} e^{-rT} \psi(T) \, \mathrm{d}T = p_0 g + e^{-rT_1^*} \chi(T_1^*). \tag{8}$$

(e) *If G has finite support $\{\lambda_1, \dots, \lambda_J\}$ and π_j denotes the probability that the arm is good with rate λ_j, the DM experiments if and only if*

$$\sum_{j=1}^{J} \frac{\pi_j}{\alpha(\lambda_j)} > 1. \tag{9}$$

The stopping boundary is the hyperplane through the points at which each edge joining "bad" to "good with rate λ_j" crosses the known-rate cutoff $\alpha(\lambda_j)$.

The proof is short given Lemma 2. Since the clock cannot run faster than calendar time ($T_t \leq t$), the payoff $\int e^{-rt}\psi(T_t)\,\mathrm{d}T_t$ is bounded by the payoff of visiting every clock value as early as possible, $\int e^{-rT}\psi(T)\,\mathrm{d}T$, restricted to the clock values where ψ is positive. Because ψ is decreasing, these form the interval $[0, T_1^*)$. The bound is attained by experimenting at full intensity until T_1^*.

The stopping rule is *myopic*: the DM stops as soon as one more instant of experimentation, followed by stopping, is unprofitable. This is the "monotone case" of optimal stopping (Chow et al., 1971), and it holds for every distribution of arrival rates because the net benefit of experimentation, ψ, falls along every no-news path: in the normalized form (5), the expected arrival rate $\bar{\lambda}_T$ falls (Lemma 1(c)) and the myopic cost $s - gp_T$ rises. Hope slows the decline in p_T but never reverses the decline of the index.

By (7), the DM acts as if a good arm's arrival rate were known and equal to its current posterior mean. Because m_T declines with silence (Lemma 1(b)), the cutoff $\alpha(m_T)$ *rises*: as silence accumulates, the DM becomes less tolerant of doubts about viability, even though silence erodes her belief in viability more slowly than it would with a known arrival rate. Figure 2(a) illustrates the two moving parts: the belief p_T and the rising bar $\alpha(m_T)$.

Corollary 2 (Two arrival rates)**.** *With two arrival rates $\lambda^- < \lambda^+$, the DM experiments at (p^-, p^+) if and only if $p^-/\alpha(\lambda^-) + p^+/\alpha(\lambda^+) > 1$. The stopping boundary is the segment*

$$\ell_1 := \left\{(p^-, p^+) : \frac{p^-}{\alpha(\lambda^-)} + \frac{p^+}{\alpha(\lambda^+)} = 1\right\}$$

connecting the known-rate cutoffs $(\alpha(\lambda^-), 0)$ and $(0, \alpha(\lambda^+))$. Starting from (p_0^-, p_0^+) above ℓ_1, the DM experiments until the clock reaches the unique root T_1^* of

$$(r + \lambda^-)\, p_0^- e^{-\lambda^- T} + (r + \lambda^+)\, p_0^+ e^{-\lambda^+ T} = \frac{rs\,(1 - p_0^- - p_0^+)}{g - s}. \tag{10}$$

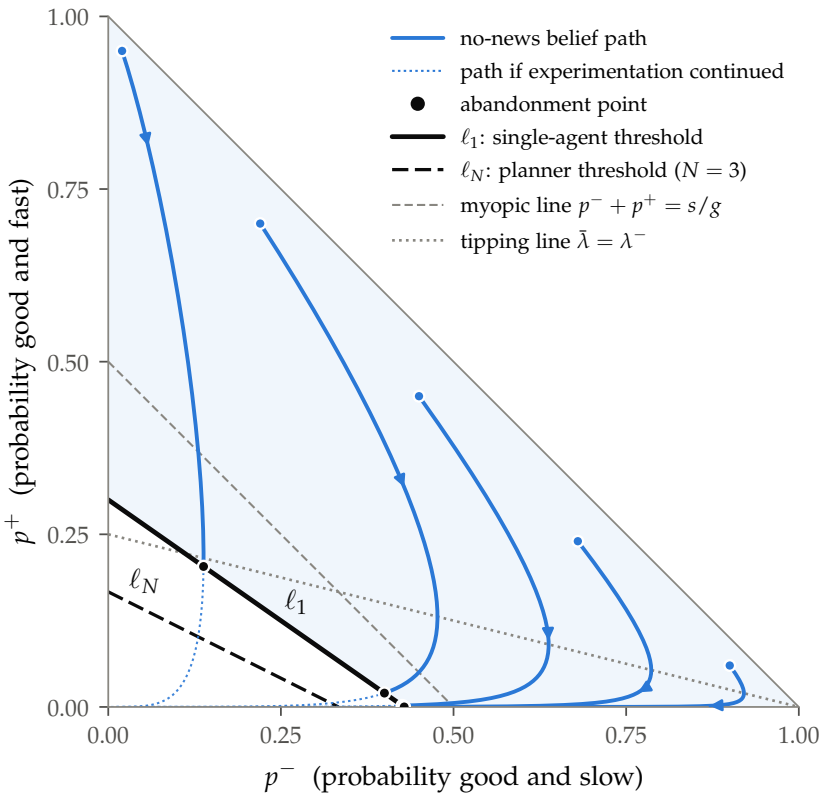


Figure 1: Beliefs and thresholds with two arrival rates. Parameters: $\lambda^- = 1$, $\lambda^+ = 4$, $g = 2$, $s = 1$, $r = 3$. Blue curves are no-news belief paths from five priors (dots). The single-agent DM stops where a path meets the segment ℓ_1 joining the known-rate cutoffs $\alpha(\lambda^-) = 3/7$ and $\alpha(\lambda^+) = 0.3$. The dashed segment ℓ_N is the efficient boundary with $N = 3$ players (Section 5). Above the dotted tipping line, the probability p^- that the arm is good but slow *rises* in the absence of news. Paths from optimistic priors first move away from the fast vertex and then run along the slow edge toward the bad vertex. Computed from the model using Python code developed with assistance from Opus 5.5 (Anthropic); see Appendix B.

Figure 1 also shows why abandonment need not reveal much about the arrival rate: paths from priors near the fast vertex reach interior points of ℓ_1 before the fast type is ruled out.

The index structure also yields the Gittins index of the risky arm. For a belief μ, let $\gamma(\mu)$ be the safe payoff that makes the DM indifferent between retiring at μ and continuing optimally, that is, the infimum of the safe payoffs s' for which the value at μ equals s'.

Corollary 3 (Index)**.** *The Gittins index of the risky arm is*

$$\gamma(\mu) = \frac{r \cdot g p + \bar{\lambda} \cdot g}{r + \bar{\lambda}}, \tag{11}$$

a weighted average of the current expected flow payoff gp and the post-breakthrough payoff g, with weights given by the discount rate r and the expected arrival rate $\bar{\lambda}$. Along every no-news path with $0 < p < 1$, the index strictly decreases.

With several independent projects of this kind, the Gittins index theorem implies that it is optimal to work on a project with the highest index (Gittins, 1979; Gittins et al., 2011; see El Karoui and Karatzas, 1994 and Kaspi and Mandelbaum, 1998 for continuous time). By (11), a project's index depends on the current belief only through the probability p that it is good and its expected arrival rate $\bar{\lambda}$, and it is increasing in both. Two projects with the same p and the same mean rate conditional on viability therefore start with the same index, however dispersed their arrival rates; how their indices evolve as they are tried depends on the dispersion, as the next section shows.

4 The economics of hope

How does rate–size uncertainty change experimentation? I compare the DM's problem under G with a *known-rate benchmark* in which every good arm has the mean arrival rate $\hat{\lambda}$ of G, and hence lump sums of mean $g/\hat{\lambda}$, keeping p_0, g, s, and r fixed. I use the subscripts G and δ for the two problems. More generally, G' is *more dispersed* than G, written $G \leq_{\mathrm{cx}} G'$, if $\int \phi \, \mathrm{d}G \leq \int \phi \, \mathrm{d}G'$ for every convex ϕ (the convex order; Shaked and Shanthikumar, 2007). The benchmark $\delta_{\hat{\lambda}}$ is the least dispersed distribution with mean $\hat{\lambda}$.

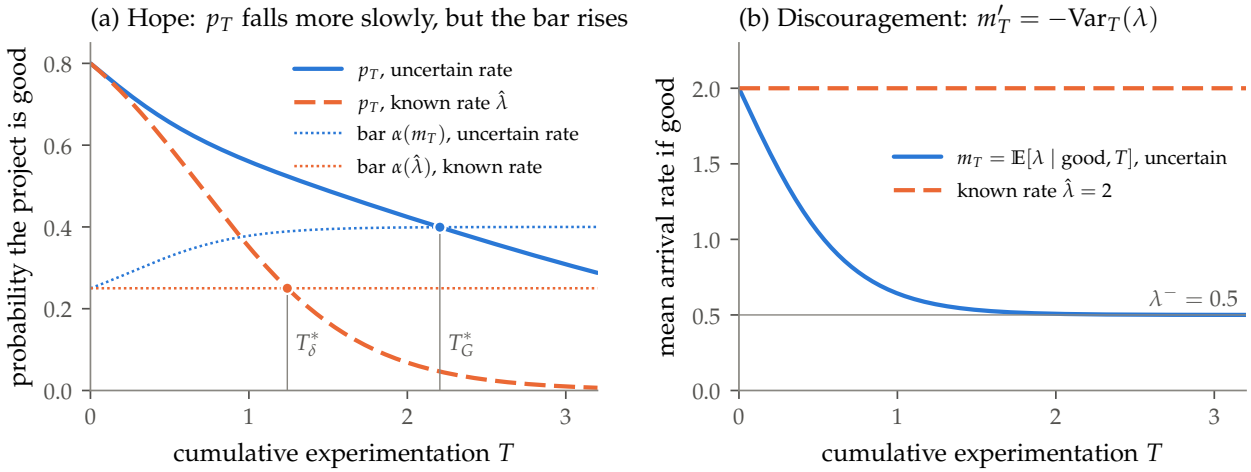


Figure 2: Hope and discouragement along a no-news path. Parameters: $g = 2$, $s = 1$, $r = 1$, $p_0 = 0.8$. Uncertain rate: $\lambda \in \{0.5, 3.5\}$ with equal probabilities; known rate: $\hat{\lambda} = 2$. Panel (a): the probability that the arm is good falls more slowly under rate uncertainty (hope), but the bar $\alpha(m_T)$ that it must clear rises because the conditional mean rate m_T falls (panel (b), discouragement). The DM stops where p_T meets the bar: at $T^*_G \approx 2.21$ with $p \approx 0.40$ under uncertainty, and at $T^*_\delta \approx 1.24$ with $p = 0.25$ under the known-rate benchmark. Computed from the model using Python code developed with assistance from Opus 5.5 (Anthropic); see Appendix B.

Proposition 1 (Unambiguous effects of rate–size uncertainty)**.** *Let G be non-degenerate with mean $\hat{\lambda}$.*

(a) (Starting.) *The DM starts experimenting under G if and only if she does under the known-rate benchmark, namely if and only if $p_0 > \alpha(\hat{\lambda})$.*

(b) (Value.) *$V_{1,G} \leq V_{1,\delta}$, with strict inequality if $p_0 > \alpha(\hat{\lambda})$. More generally, $G \leq_{\text{cx}} G'$ implies $V_{1,G'} \leq V_{1,G}$, with strict inequality if $G' \neq G$ and $p_0 > \alpha(\hat{\lambda})$.*

Suppose $p_0 > \alpha(\hat{\lambda})$. Then:

(c) (Hopeful abandonment.) *When the DM abandons the risky arm, the probability that it is good is*

$$p_{T^*_G} = \alpha\left(m_{T^*_G}\right) > \alpha(\hat{\lambda}) = p^{\delta}_{T^*_\delta}.$$

(d) (Lost good projects.) *The probability that a good arm is abandoned is*

$$L(T^*_G) = \frac{rs(1-p_0)}{(g-s)p_0} \cdot \frac{1}{r + m_{T^*_G}} > \frac{rs(1-p_0)}{(g-s)p_0} \cdot \frac{1}{r+\hat{\lambda}} = e^{-\hat{\lambda} T^*_\delta}.$$

Part (a) holds because, at the outset, the index (7) depends on the distribution of arrival rates only through its mean. Part (b) holds because, for any fixed no-news stopping clock $T > 0$, the uncertainty lowers the probability of a breakthrough before stopping:

by Jensen's inequality, $L(T) = \mathbb{E}[e^{-\lambda T}] > e^{-\hat{\lambda} T}$. Parts (c) and (d) express hope and discouragement at the moment of abandonment: the DM gives up on projects she still considers relatively likely to be good, because by then she believes that a good project would be slow. In the example of Fig. 2, she abandons when the probability that the arm is good is 0.40, compared with 0.25 under the known-rate benchmark, and she abandons a good arm with probability 0.17 rather than 0.08; the value falls from 1.64 to 1.61.

Parts (c) and (d) hold whether the uncertainty makes the DM more or less persistent in the absence of a breakthrough; which of the two occurs is the subject of the next theorem. Persistence is measured by the *no-news abandonment horizon* T^*, the amount of experimentation the DM is willing to undertake without a breakthrough; because the optimal policy experiments at full intensity until success or abandonment, T^* is also the calendar date at which the DM abandons the arm if no breakthrough occurs. Under the known-rate benchmark, the horizon is $T^*_\delta = \hat{\lambda}^{-1} \ln\big[(g-s)p_0(r+\hat{\lambda})/(rs(1-p_0))\big]$, which increases continuously from 0 to ∞ as p_0 rises from $\alpha(\hat{\lambda})$ to 1.

Theorem 2 (Abandonment horizon: hope versus discouragement)**.** *Let G be non-degenerate with mean $\hat{\lambda}$. There is a unique $\hat{T} \in (0, 2/r)$ such that, for every prior $p_0 > \alpha(\hat{\lambda})$,*

$$T^*_G < T^*_\delta \ \text{ if } T^*_\delta < \hat{T}, \qquad T^*_G = T^*_\delta \ \text{ if } T^*_\delta = \hat{T}, \qquad T^*_G > T^*_\delta \ \text{ if } T^*_\delta > \hat{T}.$$

Equivalently, with $\hat{p}_0 := rs/\big[rs + (g-s)(r+\hat{\lambda})e^{-\hat{\lambda}\hat{T}}\big] \in (\alpha(\hat{\lambda}), 1)$, rate–size uncertainty shortens the abandonment horizon if $\alpha(\hat{\lambda}) < p_0 < \hat{p}_0$ and lengthens it if $p_0 > \hat{p}_0$. If G is supported in $(0, \lambda_{\max}]$, then $\hat{T} > 2/(r + \lambda_{\max})$.

Rate–size uncertainty therefore *shortens the no-news abandonment horizon on less promising projects and lengthens it on more promising ones*. The proof compares the indices directly. Let $F_G(T) := \int (r+\lambda) e^{-\lambda T} G(\mathrm{d}\lambda)$, so that the DM stops when $(g-s)p_0 F_G(T) = rs(1-p_0)$. The ratio $R(T) := F_G(T)/F_\delta(T)$ starts at $R(0) = 1$ with slope $R'(0) = -\mathrm{Var}_G(\lambda)/(r + \hat{\lambda}) < 0$ (or $-\infty$ if the variance is infinite), is strictly convex, and grows without bound, so it crosses 1 exactly once, at $\hat{T}$, from below.

The negative initial slope is discouragement: early silence quickly rules out the fast types, whose imminent breakthroughs account for much of the initial value of experimenting, so the index falls faster than under the known-rate benchmark. The convexity is hope: slow types are hard to rule out, and eventually the lowest arrival rates govern the decay of the index, just as the far future is discounted at the lowest rate in Weitzman (1998). A less promising project is abandoned early, before hope can take hold; a more promising one is pursued long enough for hope to dominate. The bound $\hat{T} < 2/r$

gives a simple sufficient condition: if the known-rate abandonment horizon exceeds $2/r$, twice the discounting horizon, then every non-degenerate distribution of arrival rates with the same mean lengthens the abandonment horizon.

The same crossing orders Gittins indices. Along a no-news path, the index (11) is an increasing function of $F_G(T)$ (see the proof of Corollary 3). Consider two projects with the same p_0 and mean arrival rate, one with an uncertain rate and one with a known arrival rate. After the same amount of experimentation T, the project with an uncertain rate has the lower index if $T < \hat{T}$ and the higher index if $T > \hat{T}$. Since a DM who works on both under the index policy keeps their indices equal, she initially experiments less with the project with an uncertain rate, but more if experimentation goes on long enough.

Remark 2 (Abandonment horizons and observed durations). Let B be the first breakthrough time under full experimentation, with $B = \infty$ for a bad arm, and let $\tau = \min\{B, T^*\}$ be the time until success or abandonment. Since $\Pr(B > t) = D(t)$,

$$\mathbb{E}[\tau] = \int_0^{T^*} D(t)\,\mathrm{d}t = (1 - p_0)T^* + p_0 \int_0^{T^*} L(t)\,\mathrm{d}t.$$

The ordering of abandonment horizons in Theorem 2 need not carry over to $\mathbb{E}[\tau]$. For example, let $r = s = 1$, $g = 2$, $p_0 = 0.58$, and $G = \frac{1}{2}\delta_{0.5} + \frac{1}{2}\delta_{3.5}$, with known-rate benchmark δ_2. Then $T^*_G \approx 0.698 < T^*_\delta \approx 0.711$, whereas $\mathbb{E}[\tau_G] \approx 0.540 > \mathbb{E}[\tau_\delta] \approx 0.518$. Such a reversal can occur only when uncertainty shortens the horizon: if $T^*_G \geq T^*_\delta > 0$, then $\mathbb{E}[\tau_G] > \mathbb{E}[\tau_\delta]$, because $D_G(t) > D_\delta(t)$ for every $t > 0$.

Figure 3 suggests a limiting polarization of abandonment horizons. With the mean arrival rate held fixed, extreme uncertainty means that almost all good arms are extremely slow, paying very large sums very rarely, while a vanishing fraction are extremely fast.

Proposition 2 (Extreme rate–size uncertainty: polarization)**.** *Let $(G_i)_{i\geq 1}$ be distributions of arrival rates with common mean $\hat{\lambda}$ such that $G_i((\varepsilon, \infty)) \to 0$ for every $\varepsilon > 0$, and fix $p_0 \in (0,1)$. As $i \to \infty$:*

(a) if $gp_0 > s$, then $T^*_{G_i} \to \infty$; if $gp_0 < s$, then $T^*_{G_i} \to 0$;

(b) $V_{1,G_i} \to \max\{gp_0, s\}$: the value of information vanishes.

The few fast types are ruled out almost instantly, while silence is nearly uninformative about the many slow ones, leaving little value to learning. At the knife edge $gp_0 = s$, the proposition establishes only that the value converges to s; it asserts no limit for the abandonment horizon. The hypothesis holds for gamma-distributed arrival rates with fixed mean and shape parameter tending to zero, as in Fig. 3, and for two-point

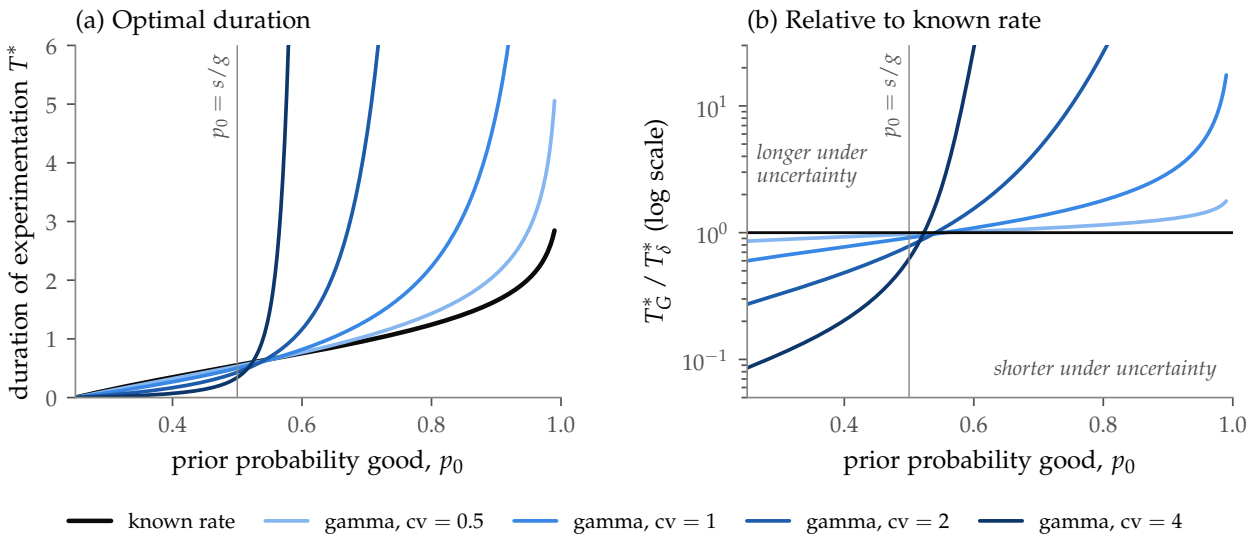


Figure 3: No-news abandonment horizons and rate–size uncertainty. Parameters: $g = 2$, $s = 1$, $r = 1$, mean arrival rate $\hat{\lambda} = 2$, so that $\alpha(\hat{\lambda}) = 0.25$ and $s/g = 0.5$. Uncertain arrival rates are gamma distributed with coefficient of variation cv $\in \{0.5, 1, 2, 4\}$. Panel (a): the abandonment horizon T^* as a function of the prior. Panel (b): that horizon relative to the known-rate benchmark. Every curve crosses 1 once, at $\hat{p}_0 \in \{0.558, 0.554, 0.541, 0.523\}$ (Theorem 2). As uncertainty grows, abandonment horizons polarize around the myopic threshold $p_0 = s/g$ (Proposition 2). Computed from the model using Python code developed with assistance from Opus 5.5 (Anthropic); see Appendix B.

distributions that put probability $(\hat{\lambda} - \varepsilon_i)/(\lambda_i - \varepsilon_i)$ on a rate $\lambda_i \to \infty$ and the rest on a rate $\varepsilon_i \to 0$.

Example 1 (Gamma-distributed arrival rates)**.** Let G be gamma with shape $a > 0$ and rate parameter $\beta > 0$, so that $\hat{\lambda} = a/\beta$ and the coefficient of variation is $a^{-1/2}$. Then $L(T) = (1 + T/\beta)^{-a}$, and

$$\frac{p_T}{1 - p_T} = \frac{p_0}{1 - p_0}\left(1 + \frac{T}{\beta}\right)^{-a}, \qquad m_T = \frac{a}{\beta + T}.$$

The odds that the arm is good decay *polynomially* rather than exponentially, and the mean arrival rate conditional on viability declines *hyperbolically*. If $p_0 > \alpha(\hat{\lambda})$, the DM stops at the unique positive root of $(g - s)p_0\big(r + a/(\beta + T)\big)(1 + T/\beta)^{-a} = rs(1 - p_0)$; otherwise, she stops immediately. As $a \to \infty$ with $a/\beta = \hat{\lambda}$, the known-rate model obtains; as $a \to 0$, Proposition 2 applies.

Table 1 collects the comparative statics.

Table 1: Effects of rate–size uncertainty, holding fixed p_0 and the mean arrival rate $\hat{\lambda}$ (rows 2–4: when $p_0 > \alpha(\hat{\lambda})$)

Outcome	Effect of the uncertainty	Result
Decision to start experimenting	none	Proposition 1(a)
Value of the problem	lower; decreasing in the convex order	Proposition 1(b)
Belief in viability at abandonment	higher (hopeful abandonment)	Proposition 1(c)
Probability a good arm is abandoned	higher	Proposition 1(d)
No-news abandonment horizon	shorter if $\alpha(\hat{\lambda}) < p_0 < \hat{p}_0$; longer if $p_0 > \hat{p}_0$	Theorem 2
Extreme uncertainty	horizons polarize around $p_0 = s/g$; information becomes worthless	Proposition 2
Teams of N players, symmetric equilibrium		
Long-run outcomes	as for a single agent	Proposition 4(a)
Proportional loss of good projects due to free riding	lower when $p_0 > \alpha(\hat{\lambda})$	Proposition 4(b)
Welfare cost of free riding	ambiguous; vanishes in the limit	Proposition 4(c), Fig. 5

5 Strategic experimentation

5.1 The game

There are $N \geq 2$ players, each with one unit of resource per unit of time and a replica of the two-armed bandit of Section 2. The safe arms are identical, and the risky arms are of a *common* type: either all are bad, or all are good with a common arrival rate Λ, drawn as in Assumption 1. Conditional on Λ, lump sums on player n's risky arm arrive according to a Poisson process with intensity $k_{n,t}\Lambda$, independently across players. Actions and outcomes are publicly observed, so players share a common posterior, and each player maximizes her own expected payoff as in Section 2. With $G = \delta_\lambda$ this is the game of Keller et al. (2005).

Let $K_t := \sum_n k_{n,t}$ be the aggregate intensity. Before the first breakthrough, the common posterior is indexed by the aggregate clock $T_t := \int_0^t K_u \, \mathrm{d}u$; after a breakthrough, using the risky arm forever is dominant for every player. Following Keller et al. (2005), I restrict players to Markov strategies with the common posterior as the state variable. Since the no-news posteriors form the one-parameter family indexed by the clock, a Markov strategy is described by a function $k_n : [0, \infty) \to [0, 1]$ of the clock. In the two-rate model,

with the triangle as the state space, this is the restriction of a strategy to the no-news curve through the prior.

A strategy is *admissible* if it is right-continuous and piecewise Lipschitz: every compact interval can be partitioned into finitely many intervals on each of which k_n is Lipschitz continuous. Since beliefs move "downward" as the clock runs, right-continuity in the clock corresponds to the left-continuity in beliefs required by Keller et al. (2005). Given admissible strategies, the clock is the unique solution of $\mathrm{d}T_t/\mathrm{d}t = K(T_t)$. Given admissible strategies of the others and any measurable deviation $k_n(t)$ of player n, it is the unique solution of $\mathrm{d}T_t/\mathrm{d}t = k_n(t) + K_{-n}(T_t)$, where $K_{-n} := \sum_{m \neq n} k_m$. A *Markov perfect equilibrium* (MPE) is a profile of admissible strategies such that, from every state, each player's strategy is a best response to the others' strategies among all measurable deviations.

5.2 Experimentation as a public good

The payoff representation of Lemma 2 extends to the game and reveals its public-good structure. For a player and a clock value T, let $Z_n(T) := D(T)\big(u_n(T) - s\big)$, where $u_n(T)$ is the player's continuation payoff at the posterior μ_T. In words, Z_n is the player's surplus over the safe arm, weighted by the probability of reaching T in silence.

Lemma 3 (Experimentation as a public good)**.** *Fix admissible strategies of the other players and a measurable policy $k_n(t)$ of player n, and let T_t be the resulting clock started at $T_0 = T$. Then*

$$Z_n(T) = \int_0^\infty e^{-rt}\Big[\chi'(T_t)\,K_{-n}(T_t) + \psi(T_t)\,k_n(t)\Big]\,\mathrm{d}t. \tag{12}$$

Consequently, the players' average surplus is $\int_0^\infty e^{-rt}\psi_N(T_t)\,\mathrm{d}T_t$, where $\psi_N := \chi' - (r/N)\chi$.

Each unit of experimentation, by anybody, produces the public good $\chi'(T)\,\mathrm{d}T$: the gain $g - s$ times the ex ante probability that this unit yields the first breakthrough. A player's own experimentation also costs her the private flow loss $r\chi$, so her net private return is $\psi = \chi' - r\chi$, exactly the single-agent index, whereas experimentation by others is a free contribution. The game is thus one of dynamic contributions to a public good (Admati and Perry, 1991; Marx and Matthews, 2000), in which contributions accumulate into the clock and the Laplace transform of the distribution of arrival rates is the "technology" that converts cumulative contributions into the public good. The welfare-relevant index $\psi_N = \psi + r(1 - 1/N)\chi$ exceeds the private index wherever experimentation is myopically costly ($\chi > 0$); this wedge is the free-rider problem.

5.3 Efficient experimentation

Proposition 3 (Efficient experimentation)**.** *A planner who maximizes the players' average payoff has all players experiment until a breakthrough occurs or the clock reaches* $T_N^* := \inf\{T \geq 0 : \psi_N(T) \leq 0\}$*, and, if no breakthrough has occurred, has all use the safe arm thereafter. Equivalently, at every belief all players experiment if and only if*

$$(g-s)\,N\bar{\lambda} > r\,(s-gp) \qquad \Longleftrightarrow \qquad p > \alpha(Nm).$$

With finite support, the efficient boundary is the hyperplane $\sum_j \pi_j/\alpha(N\lambda_j) = 1$*; with two arrival rates, it is the segment* ℓ_N *joining* $(\alpha(N\lambda^-), 0)$ *and* $(0, \alpha(N\lambda^+))$*. If* $T_1^* > 0$*, then* $T_N^* > T_1^*$*. The average value at the prior is* $V_N^P = p_0 g + e^{-rT_N^*/N}\chi(T_N^*)$*.*

The planner is a single DM whose arm has its arrival rate multiplied by N. Equivalently, measured on the clock, she is a single DM with discount rate r/N. All the results of Sections 3 and 4 therefore apply to efficient experimentation, with r replaced by r/N.

5.4 The symmetric equilibrium

Theorem 3 (Symmetric equilibrium)**.** *Suppose* $p_0 > \alpha(\hat{\lambda})$*, so that* $T_1^* > 0$*. The game has a unique symmetric MPE. Along the no-news path, each player's intensity is*

$$k^\dagger(T) = \begin{cases} 1, & 0 \leq T < T^\dagger, \\ \dfrac{1}{N-1} \cdot \dfrac{\int_T^{T_1^*} \psi(\tau)\,\mathrm{d}\tau}{\chi(T)}, & T^\dagger \leq T < T_1^*, \\ 0, & T \geq T_1^*, \end{cases} \tag{13}$$

where $T^\dagger \in [0, T_1^*)$ *is the unique root of* $\int_T^{T_1^*} \psi(\tau)\,\mathrm{d}\tau = (N-1)\chi(T)$ *if* $\int_0^{T_1^*} \psi > (N-1)\chi(0)$*, and* $T^\dagger = 0$ *otherwise. Moreover:*

(a) $k^\dagger$ *is continuous and strictly decreasing on* $[T^\dagger, T_1^*]$*, with* $k^\dagger(T^\dagger) = 1$ *if* $T^\dagger > 0$ *and* $k^\dagger(T) \to 0$ *as* $T \uparrow T_1^*$*. Players mix only where experimentation is myopically unprofitable* ($p_T < s/g$)*.*

(b) The equilibrium clock converges to T_1^* *but never reaches it: beliefs converge to the single-agent boundary without ever reaching it.*

(c) Each player's equilibrium value at the prior is

$$V_N^E = \begin{cases} s + N\chi(T^\dagger)e^{-rT^\dagger/N} - \chi(0), & T^\dagger > 0, \\ s + \int_0^{T_1^*} \psi(\tau)\,\mathrm{d}\tau, & T^\dagger = 0, \end{cases} \qquad \textit{and} \qquad V_1 < V_N^E < V_N^P.$$

If $p_0 \le \alpha(\hat{\lambda})$, the unique symmetric MPE has all players use the safe arm forever.

The structure is that of Keller et al. (2005), to which Theorem 3 reduces when G is degenerate. While the arm is myopically attractive, and for a while after, all players experiment. As the arm becomes less promising, each player would rather that the others experiment, and the players mix at the intensity $k^\dagger$, which keeps each of them indifferent between experimenting and free riding at every instant; in this mixing region, each player's surplus $Z(T) = \int_T^{T_1^*} \psi$ does not depend on N. What is new is that (13) holds for *every* distribution of arrival rates, despite the multidimensional state, and that each of its ingredients is explicit given the Laplace transform L.

The public-good structure reduces the uniqueness proof to a scalar terminal-value problem with a Lipschitz right side.

With two arrival rates, the boundary $\mathcal{C}^\dagger$ between full experimentation and mixing lies between ℓ_1 and the myopic line $p^- + p^+ = s/g$ (Fig. 4(a)).

Corollary 4 (Too little, too slowly)**.** *Suppose $p_0 > \alpha(\hat{\lambda})$. In the symmetric equilibrium, experimentation stops (asymptotically) at the single-agent boundary rather than the efficient one, since $T_1^* < T_N^*$, and the aggregate intensity is below the efficient intensity N on $(T^\dagger, T_1^*)$. Near the boundary the aggregate intensity $Nk^\dagger(T)$ falls below 1: a team of N players experiments less intensely than a lone DM would.*

In Fig. 4(b), the team initially learns faster than a single agent, but once players mix, experimentation slows and the single agent's belief path overtakes the equilibrium path. The equilibrium value is 1.529, against 1.509 for a single agent and 1.548 under the efficient policy.

5.5 Free riding meets hope

Rate–size uncertainty also changes the severity of the free-rider problem.

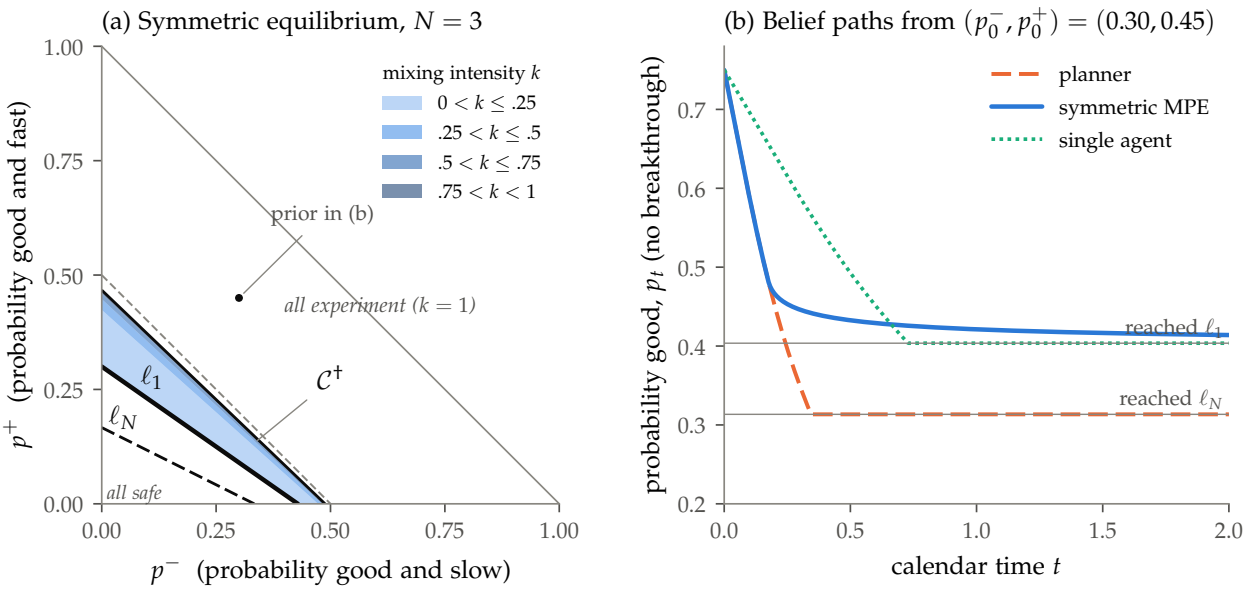


Figure 4: The symmetric equilibrium with two arrival rates. Parameters: $\lambda^- = 1$, $\lambda^+ = 4$, $g = 2$, $s = 1$, $r = 3$, $N = 3$. Panel (a): all players experiment above $\mathcal{C}^\dagger$; they mix with intensity $k^\dagger$ (shading) between $\mathcal{C}^\dagger$ and the single-agent boundary ℓ_1; all play safe below ℓ_1. The efficient boundary is ℓ_N, and the dashed gray line is the myopic line. Panel (b): calendar-time paths of p_t from the prior $(0.30, 0.45)$. The planner reaches ℓ_N at $t \approx 0.34$ and the single agent reaches ℓ_1 at $t \approx 0.73$; the equilibrium path approaches ℓ_1 only asymptotically. Computed from the model using Python code developed with assistance from Opus 5.5 (Anthropic); see Appendix B.

Proposition 4 (Free riding and rate–size uncertainty)**.** *Let G be non-degenerate with mean $\hat{\lambda}$. Parts (a) and (b) assume $p_0 > \alpha(\hat{\lambda})$.*

(a) (Long-run outcomes.) *In the symmetric equilibrium, the total amount of experimentation (the limit of the clock) is T_1^*, a good arm is never developed with probability $L(T_1^*)$, and experimentation dies out at the belief $p_{T_1^*} = \alpha(m_{T_1^*})$. Hence Proposition 1(c)–(d) and Theorem 2 apply verbatim to equilibrium play, with Theorem 2 comparing total amounts of experimentation.*

(b) (Proportional loss.) *Free riding raises the probability that a good project is never developed by the factor*

$$\frac{L(T_1^*)}{L(T_N^*)} = \frac{r + N\, m_{T_N^*}}{r + m_{T_1^*}} \;<\; \frac{r + N\hat{\lambda}}{r + \hat{\lambda}},$$

where the right side is the same factor under the known-rate benchmark.

(c) (Extreme uncertainty.) *For any sequence (G_i) as in Proposition 2 and any $p_0 \in (0,1)$, the single-agent, equilibrium, and efficient values all converge to $\max\{gp_0, s\}$. In particular, the welfare cost of free riding, $V_N^P - V_N^E$, converges to zero.*

Parts (b) and (c) reflect the declining informativeness of silence as the posterior concentrates on low arrival rates: the additional experimentation undertaken by the planner becomes less informative. Under the polarization hypothesis, silence becomes uninformative about the low arrival rates that dominate the prior, and the free-rider problem disappears together with the value of information.

For intermediate uncertainty, the welfare cost of free riding can rise or fall (Fig. 5); the examples give no general criterion for the sign. In all the numerical comparisons considered, the equilibrium value V_N^E itself decreases with dispersion in arrival rates. The corresponding monotonicity results for the single-agent and efficient values follow from Proposition 1(b) and Proposition 3; whether the equilibrium value decreases under arbitrary mean-preserving spreads remains open.

6 Extensions and discussion

6.1 Rate-dependent flow payoffs

The optimal-policy and equilibrium characterizations extend to flow payoffs that depend on the arrival rate when every viable type remains preferable to the safe arm.

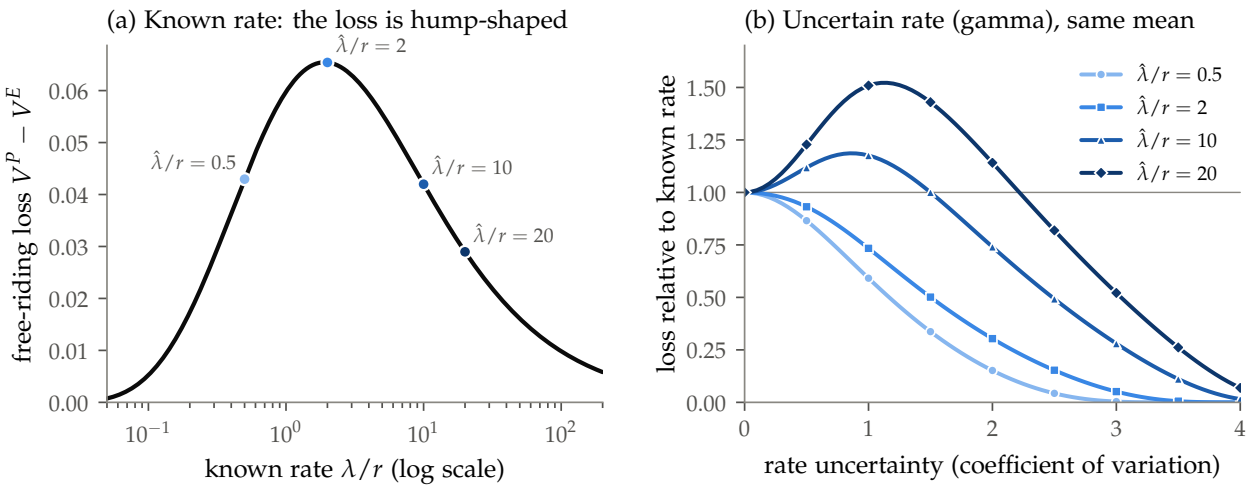


Figure 5: The welfare cost of free riding, $V_N^P - V_N^E$, evaluated at the prior. Parameters: $g = 2$, $s = 1$, $r = 1$, $N = 3$, $p_0 = 0.6$. Panel (a): for these parameters, the known-rate cost is hump-shaped in λ/r, peaking near $\lambda/r \approx 1.9$. Panel (b): the cost with gamma-distributed arrival rates of mean $\hat{\lambda}$, relative to the known-rate cost, as a function of the coefficient of variation of arrival rates. In the plotted comparisons, the uncertainty lowers the cost at $\hat{\lambda}/r \in \{0.5, 2\}$. At $\hat{\lambda}/r \in \{10, 20\}$, moderate uncertainty raises it by up to 19% and 52%, respectively. In the polarization limit, the cost converges to zero (Proposition 4(c)). Computed from the model using Python code developed with assistance from Opus 5.5 (Anthropic); see Appendix B.

Proposition 5 (Rate-dependent flow payoffs)**.** *Suppose that a good arm with rate λ yields lump sums of mean $g(\lambda)/\lambda$, where $g(\lambda) > s$ for G-almost every λ and $\int (1+\lambda)(g(\lambda) - s)\, G(\mathrm{d}\lambda) < \infty$. Let $c := \int (g(\lambda) - s)\, G(\mathrm{d}\lambda)$. Then, for every profile of no-news policies, with each player using the risky arm forever after the first breakthrough, every player's ex ante payoff coincides with her payoff under the same policies in the model of Assumption 1 with flow payoff $\tilde{g} := s + c$ and the payoff-weighted rate distribution $\tilde{G}(\mathrm{d}\lambda) := (g(\lambda) - s)\, G(\mathrm{d}\lambda)/c$. Consequently, Theorems 1 and 3 and Proposition 3 hold with (g, G) replaced by $(\tilde{g}, \tilde{G})$. With finitely many arrival rates, the single-agent boundary is the hyperplane $\sum_j \pi_j/\alpha_j = 1$ with type-specific cutoffs $\alpha_j := rs/\big(rg(\lambda_j) + \lambda_j(g(\lambda_j) - s)\big)$.*

The equivalence concerns payoffs under the same no-news policies; Bayesian beliefs still use G. The comparative statics apply only when their hypotheses hold for $(\tilde{g}, \tilde{G})$: a mean-preserving spread of G need not induce one of $\tilde{G}$ and may change $\tilde{g}$, whereas the earlier comparisons fix the payoff and mean arrival rate.

When some viable types yield less than the safe arm, so that $g(\lambda) < s$ for some λ, a breakthrough no longer settles whether to continue: it reveals that the arm is good, but the DM must keep learning about its rate, as in the Poisson bandit of Keller and Rady (2010), and she may eventually abandon an arm that has paid off. In the game, this

post-breakthrough learning creates an encouragement effect that is absent here. I leave this case to future work.

6.2 Other equilibria and information structures

The public-good formulation may also permit asymmetric MPE in which players take turns experimenting, as in Keller et al. (2005); I do not pursue them. Another variation lets the arrival rate be player-specific while viability is common. Joint silence would then be less likely to reflect common slowness: because $\mathbb{E}[e^{-\lambda NT}] \geq (\mathbb{E}[e^{-\lambda T}])^N$, each player's silence would be more informative about viability than under a common rate, and hope would be weaker.

6.3 Applications and empirical implications

Patrons and artists A publisher finances a writer who, if talented, may produce great books rarely or mediocre ones often; s is the return on the publisher's best alternative use of funds. Relative to a known pace with the same mean, Theorem 2 calls for more patience with a highly promising writer and less with a modestly promising one while awaiting a first success.

Research and development The baseline model fits viable research programs yielding either occasional major results or frequent minor ones, with the same expected value per unit of time; Proposition 5 allows that value to vary with the arrival rate while exceeding the safe payoff for every good type. The declining success hazard partly reflects dynamic selection (Heckman and Singer, 1984; Lancaster, 1990): silent projects are disproportionately bad or slow. For research teams that begin experimenting, effort along a no-news path stays positive at every finite time but tends to zero, with cumulative effort below the efficient level. Observed durations also depend on breakthrough timing (Remark 2).

One-shot prizes A first success may end the problem: a drug is approved, a patent granted, or a job found. Suppose it yields a prize $W > 0$ independent of the arrival rate, experimentation costs a flow $\kappa > 0$ per unit of intensity, and stopping yields nothing. The DM experiments if and only if $\bar{\lambda}_T W > \kappa$, a myopic rule independent of r (see the appendix derivation). If G is non-degenerate and experimentation is initially worthwhile ($p_0 \hat{\lambda} W > \kappa$), abandonment occurs after a positive amount of experimentation at a belief

$p_T = \kappa/(Wm_T) > \kappa/(W\hat{\lambda})$. With a known rate, the inequality becomes an equality whenever experimentation begins; if $p_0\hat{\lambda}W \leq \kappa$, stopping is immediate. Thus a job seeker unsure about both viability and the arrival rate of offers may search and then give up at a more optimistic belief than under the known-rate benchmark (compare Gonzalez and Shi, 2010). This establishes the stopping rule and hopeful abandonment; the baseline model's other comparative statics do not automatically carry over.

7 Conclusion

Rate–size uncertainty makes silence ambiguous: it weakens the evidence against viability while lowering the mean arrival rate conditional on viability. Relative to a known-rate benchmark with the same mean, uncertainty lowers the value whenever experimentation is worthwhile, yet lengthens the no-news abandonment horizon on more promising projects and shortens it on less promising ones. The expected time until success or abandonment need not move in the same direction.

These forces also govern cumulative experimentation in teams. Numerical examples show that uncertainty can raise or lower the welfare cost of free riding, while the polarization limit eliminates that cost. A natural next step is the design of contracts and deadlines when an agent's arrival rate is unknown to both sides.

A Proofs

Throughout, $\nu_0 \coloneqq (1-p_0)\delta_0 + p_0 G$ denotes the prior distribution of the arrival rate Λ, and ν_T the posterior after T units of experimentation without a breakthrough. Differentiation under the integral sign in expressions such as $\int \lambda^k e^{-\lambda T} G(\mathrm{d}\lambda)$ is justified for $T > 0$ by dominated convergence, since $\lambda^k e^{-\lambda T}$ is bounded on $(0, \infty)$. At $T = 0$, only first derivatives are used (as right derivatives, which exist because G has a finite mean), except where a finite variance is assumed explicitly.

Proof of Lemma 1

(a) By Bayes' rule, $\nu_T(\mathrm{d}\lambda) = e^{-\lambda T}\nu_0(\mathrm{d}\lambda)/D(T)$. Hence $\nu_T(\{0\}) = (1-p_0)/D(T)$, $p_T = p_0 L(T)/D(T)$, and conditional on $\Lambda > 0$ the rate has distribution $G_T(\mathrm{d}\lambda) = e^{-\lambda T}G(\mathrm{d}\lambda)/L(T)$ with mean $m_T = \int \lambda e^{-\lambda T} G(\mathrm{d}\lambda)/L(T) = -L'(T)/L(T)$.

(b) $\ln\big(p_T/(1-p_T)\big) = \ln\big(p_0/(1-p_0)\big) + \ln L(T)$, whose derivative is $L'(T)/L(T) = -m_T$. Moreover, for $T > 0$ (and at $T = 0$ if G has finite variance),

$$m_T' = -\frac{L''(T)L(T) - L'(T)^2}{L(T)^2} = -\Big(\int \lambda^2 G_T(\mathrm{d}\lambda) - m_T^2\Big) = -v_T.$$

(c) For a Borel set A, $\nu_T(A) = \int_A e^{-\lambda T}\nu_0(\mathrm{d}\lambda)/D(T)$. Moreover,

$$-\frac{D'(T)}{D(T)} = \frac{p_0\int \lambda e^{-\lambda T}G(\mathrm{d}\lambda)}{D(T)} = \mathbb{E}[\Lambda \mid T] = \bar\lambda_T.$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}T}\nu_T(A) = -\int_A \lambda\,\nu_T(\mathrm{d}\lambda) + \nu_T(A)\,\bar\lambda_T = \mathbb{E}\big[(\bar\lambda_T - \Lambda)\mathbf{1}\{\Lambda \in A\} \mid T\big].$$

With finitely many values, taking $A = \{\lambda_j\}$ shows that $\nu_T(\{\lambda_j\})$ rises if and only if $\lambda_j < \bar\lambda_T$. Finally, $\bar\lambda_T = -D'(T)/D(T)$, so $\bar\lambda_T' = -\big(D''D - D'^2\big)/D^2 = -\big(\mathbb{E}[\Lambda^2 \mid T] - \bar\lambda_T^2\big)$.

(d) For $T > 0$, $\lambda \mapsto e^{-\lambda T}$ is strictly convex, so Jensen's inequality gives $L(T) > e^{-\hat\lambda T}$ whenever G is non-degenerate. Since $p_T = p_0L(T)/(1 - p_0 + p_0L(T))$ is increasing in $L(T)$, $p_T > p_T^\delta$. For $T > 0$, G_T is equivalent to G and hence non-degenerate, so $v_T > 0$ and m is strictly decreasing on $(0,\infty)$. Since $m_T \to m_0 = \hat\lambda$ as $T \downarrow 0$, $m_T < \hat\lambda$ for $T > 0$. □

Proof of Corollary 1

Apply Lemma 1(c) with the three values $0 < \lambda^- < \lambda^+$. At interior beliefs, $\bar\lambda_T = \lambda^- p_T^- + \lambda^+ p_T^+ < \lambda^+$, so p_T^+ decreases, and $\bar\lambda_T > 0$, so the probability $1 - p_T$ of the bad arm increases. Moreover, p_T^- increases if and only if $\bar\lambda_T > \lambda^-$. Since $\bar\lambda_T$ is strictly decreasing (its derivative is minus a positive variance) and tends to 0, the sign of $\bar\lambda_T - \lambda^-$ changes at most once, from positive to negative. □

Proof of Lemma 2

After a breakthrough, the arm is known to be good and using R forever is optimal, with expected flow payoff g; I take this continuation as given. Let k be a policy and T_t its clock. At time t, a breakthrough has occurred with probability $1 - D(T_t)$, in which case the expected flow payoff is g. Otherwise, with probability $D(T_t)$, it is $(1 - k_t)s + k_t g p_{T_t}$. Using $D(T)(s - gp_T) = \chi(T)$, the expected flow payoff at t is $s + A(T_t) - k_t\chi(T_t)$, where $A(T) := (g - s)\big(1 - D(T)\big) = (g - s)p_0\big(1 - L(T)\big)$. The function A is bounded, $A(0) =$

0, and $A' = \chi'$. Since $t \mapsto A(T_t)$ is absolutely continuous with derivative $\chi'(T_t)k_t$, integration by parts gives

$$\begin{aligned}\int_0^\infty re^{-rt}A(T_t)\,\mathrm{d}t &= \Big[-e^{-rt}A(T_t)\Big]_0^\infty + \int_0^\infty e^{-rt}\chi'(T_t)\,k_t\,\mathrm{d}t \\ &= \int_0^\infty e^{-rt}\chi'(T_t)\,k_t\,\mathrm{d}t.\end{aligned}$$

Therefore $J(k) = s + \int_0^\infty e^{-rt}k_t\big[\chi'(T_t) - r\chi(T_t)\big]\mathrm{d}t$, which is (6). □

Proof of Theorem 1

(a) ψ is continuous on $[0,\infty)$ by dominated convergence, since $(r+\lambda)e^{-\lambda T} \le r+\lambda$ is G-integrable. For $T > 0$, $\psi'(T) = -(g-s)p_0\int \lambda(r+\lambda)e^{-\lambda T}G(\mathrm{d}\lambda) < 0$. As $T \to \infty$, $\int (r+\lambda)e^{-\lambda T}G(\mathrm{d}\lambda) \to 0$ by dominated convergence, because G has no atom at 0; hence $\psi(T) \to -r(1-p_0)s < 0$ and $T_1^* < \infty$. Finally, $\psi(0) = (g-s)p_0(r+\hat\lambda) - r(1-p_0)s > 0$ if and only if $p_0\,[rg + \hat\lambda(g-s)] > rs$, that is, $p_0 > \alpha(\hat\lambda)$.

(b) Let k be any policy, T_t its clock, and $T_\infty := \lim_{t\to\infty} T_t \in [0,\infty]$. For $T \in [0, T_\infty)$ let $\tau(T) := \inf\{t \ge 0 : T_t \ge T\}$. Since $t \mapsto T_t$ is continuous and non-decreasing, the change-of-variables formula for Lebesgue–Stieltjes integrals (Revuz and Yor, 1999, Chapter 0, Proposition 4.9) and $T_{\tau(T)} = T$ give

$$J(k) - s = \int_0^\infty e^{-rt}\psi(T_t)\,\mathrm{d}T_t = \int_0^{T_\infty} e^{-r\tau(T)}\psi(T)\,\mathrm{d}T.$$

Because $k \le 1$, $T_t \le t$, so $\tau(T) \ge T$. Since $\psi > 0$ on $[0, T_1^*)$ and $\psi \le 0$ on $[T_1^*, \infty)$,

$$J(k) - s \le \int_0^{\min\{T_\infty, T_1^*\}} e^{-r\tau(T)}\psi(T)\,\mathrm{d}T \le \int_0^{T_1^*} e^{-rT}\psi(T)\,\mathrm{d}T.$$

The policy $k_t = \mathbf{1}\{t < T_1^*\}$ has $T_t = \min\{t, T_1^*\}$ and attains the bound. Equality in the second inequality requires $\tau(T) = T$ for almost every $T < T_1^*$, hence $T_t = t$ on $[0, T_1^*]$, that is, $k_t = 1$ for almost every $t < T_1^*$. Equality in the first requires that $\int_{T_1^*}^{T_\infty} e^{-r\tau(T)}\psi(T)\,\mathrm{d}T = 0$, hence $T_\infty \le T_1^*$, since $\psi < 0$ on (T_1^*, ∞). Thus every optimal policy coincides with $\mathbf{1}\{t < T_1^*\}$ almost everywhere.

(c) By (5), $\psi(T) > 0$ if and only if $(g-s)\bar\lambda_T > r(s - gp_T)$. Since the continuation problem at any posterior μ is the original problem with prior μ, part (b) applied at μ shows that the DM experiments at μ if and only if the inequality holds there. Writing $\bar\lambda = pm$, the inequality is $p\,[rg + (g-s)m] > rs$, that is, $p > \alpha(m)$.

(d) Since $\mathrm{d}\big(e^{-rT}\chi(T)\big)/\mathrm{d}T = e^{-rT}\psi(T)$, $V_1 = s + e^{-rT_1^*}\chi(T_1^*) - \chi(0)$, and $\chi(0) = s - gp_0$.

(e) With finite support, $p = \sum_j \pi_j$ and $\bar{\lambda} = \sum_j \pi_j \lambda_j$, so the inequality in (c) reads $\sum_j \pi_j \big[rg + \lambda_j(g-s)\big] > rs$, which is (9). On the edge where only "bad" and "good with rate λ_j" have positive probability, the boundary is $\pi_j = \alpha(\lambda_j)$. □

Proof of Corollary 2

The first statement is Theorem 1(e) with $J = 2$. With $p_0 G = p_0^- \delta_{\lambda^-} + p_0^+ \delta_{\lambda^+}$, the equation $\psi(T) = 0$ is (10), whose solution is unique because ψ is strictly decreasing. □

Proof of Corollary 3

Write $V_1(\mu; s')$ for the value when the safe payoff is s'. By Theorem 1(a) and (c) applied at μ, the DM experiments at μ if and only if $(g - s')\bar{\lambda} > r(s' - gp)$, that is, if and only if $s' < \gamma(\mu)$. If $s' \geq \gamma(\mu)$, then $V_1(\mu; s') = s'$. If $s' < \gamma(\mu)$, then $V_1(\mu; s') > s'$, because $\psi > 0$ near 0. Hence $\gamma(\mu) = \inf\{s' : V_1(\mu; s') = s'\}$. Along a no-news path, $1 - p_T = (1 - p_0)/D(T)$ and $\bar{\lambda}_T = p_0 \int \lambda e^{-\lambda T} G(\mathrm{d}\lambda)/D(T)$, so

$$\begin{aligned}\gamma(\mu_T) &= g\left[1 - \frac{r(1-p_T)}{r + \bar{\lambda}_T}\right] \\ &= g\left[1 - \frac{r(1-p_0)}{r(1-p_0) + p_0 \int (r+\lambda) e^{-\lambda T} G(\mathrm{d}\lambda)}\right],\end{aligned}$$

which is strictly decreasing in T. □

Proof of Proposition 1

(a) By Theorem 1(a), $T_G^* > 0$ if and only if $p_0 > \alpha(\hat{\lambda})$, and the same holds for $\delta_{\hat{\lambda}}$.

(b) By Theorem 1(b), $V_{1,G} = \max_{\tau \geq 0} W_G(\tau)$, where $W_G(\tau) := p_0 g + e^{-r\tau}\big[(1-p_0)s - (g-s)p_0 L_G(\tau)\big]$ is the payoff of experimenting until the clock reaches τ. If $G \leq_{\mathrm{cx}} G'$, then $L_{G'}(\tau) \geq L_G(\tau)$ for every τ, because $\lambda \mapsto e^{-\lambda\tau}$ is convex; hence $W_{G'} \leq W_G$ and $V_{1,G'} \leq V_{1,G}$. If moreover $G' \neq G$, then $L_{G'}(\tau) > L_G(\tau)$ for every $\tau > 0$. Indeed, by Strassen's theorem (Shaked and Shanthikumar, 2007, Theorem 3.A.4) there are random variables $X \sim G$ and $X' \sim G'$ with $\mathbb{E}[X' \mid X] = X$, and for the strictly convex function $\phi(\lambda) = e^{-\lambda\tau}$, $\mathbb{E}[\phi(X')] = \mathbb{E}\big[\mathbb{E}[\phi(X') \mid X]\big] \geq \mathbb{E}[\phi(X)]$, with equality only if $X' = X$ almost surely.

Since $p_0 > \alpha(\hat{\lambda})$ implies $T^*_{G'} > 0$, it follows that $V_{1,G'} = W_{G'}(T^*_{G'}) < W_G(T^*_{G'}) \le V_{1,G}$. The comparison with $\delta_{\hat{\lambda}}$ is the special case $G = \delta_{\hat{\lambda}}$, $G' = G$.

(c) $\psi(T^*_G) = 0$ and (5) give $(g-s)p\,m = r(s-gp)$ at T^*_G, that is, $p_{T^*_G} = \alpha(m_{T^*_G})$. By Lemma 1(d), $m_{T^*_G} < \hat{\lambda}$, and α is strictly decreasing.

(d) $\psi(T^*_G) = 0$ is equivalent to $(g-s)p_0 L(T^*_G)\big(r + m_{T^*_G}\big) = r(1-p_0)s$, since $-L' = mL$. Solve for $L(T^*_G)$ and use $m_{T^*_G} < \hat{\lambda}$. Under the known-rate benchmark, the same computation gives $e^{-\hat{\lambda}T^*_\delta} = rs(1-p_0)/\big[(g-s)p_0(r+\hat{\lambda})\big]$. □

Proof of Theorem 2

Let $F_G(T) := \int (r+\lambda)e^{-\lambda T}G(\mathrm{d}\lambda)$ and $F_\delta(T) := (r+\hat{\lambda})e^{-\hat{\lambda}T}$. Both are continuous and strictly decreasing, and for $p_0 > \alpha(\hat{\lambda})$ the abandonment horizons T^*_G and T^*_δ are the unique solutions of $(g-s)p_0F(T) = r(1-p_0)s$. Let

$$R(T) := \frac{F_G(T)}{F_\delta(T)} = \frac{1}{r+\hat{\lambda}} \int (r+\lambda)\, e^{-(\lambda-\hat{\lambda})T}\, G(\mathrm{d}\lambda).$$

Step 1. R is continuous on $[0,\infty)$ by dominated convergence, since $(r+\lambda)e^{-(\lambda-\hat{\lambda})T} \le (r+\lambda)e^{\hat{\lambda}T}$ is integrable, and $R(0) = 1$. For $T > 0$, $R''(T) = (r+\hat{\lambda})^{-1}\int (r+\lambda)(\lambda-\hat{\lambda})^2 e^{-(\lambda-\hat{\lambda})T}G(\mathrm{d}\lambda) > 0$, so R is strictly convex on $[0,\infty)$. Moreover, $R(T) < 1$ for all small $T > 0$. To see this, write

$$\frac{R(T)-1}{T} = \frac{1}{r+\hat{\lambda}} \int (r+\lambda)\, \frac{e^{-(\lambda-\hat{\lambda})T}-1}{T}\, G(\mathrm{d}\lambda).$$

On $\{\lambda < \hat{\lambda}\}$ the integrand converges to $(r+\lambda)(\hat{\lambda}-\lambda)$, dominated by $(r+\hat{\lambda})\hat{\lambda}e^{\hat{\lambda}}$ for $T \le 1$. On $\{\lambda > \hat{\lambda}\}$ it is non-positive and converges to $-(r+\lambda)(\lambda-\hat{\lambda})$, so by Fatou's lemma the limit superior of its integral is at most $-\int_{\{\lambda>\hat{\lambda}\}}(r+\lambda)(\lambda-\hat{\lambda})\,G(\mathrm{d}\lambda)$. Hence $\limsup_{T\downarrow 0}(R(T)-1)/T \le -(r+\hat{\lambda})^{-1}\int (r+\lambda)(\lambda-\hat{\lambda})\,G(\mathrm{d}\lambda) = -\operatorname{Var}_G(\lambda)/(r+\hat{\lambda}) < 0$, where the variance may be infinite.

Step 2. Since G is non-degenerate with mean $\hat{\lambda}$, there is $\varepsilon > 0$ with $G\big((0,\hat{\lambda}-\varepsilon)\big) > 0$. Then $R(T) \ge (r+\hat{\lambda})^{-1}\, r\, G\big((0,\hat{\lambda}-\varepsilon)\big)e^{\varepsilon T} \to \infty$.

Step 3. A continuous, strictly convex function with $R(0) = 1$, $R < 1$ near 0, and $R(T) \to \infty$ takes the value 1 at exactly one $\hat{T} > 0$, and satisfies $R < 1$ on $(0,\hat{T})$ and $R > 1$ on $(\hat{T},\infty)$.

Step 4. If $T^*_\delta < \hat{T}$, then $(g-s)p_0F_G(T^*_\delta) = R(T^*_\delta)\,r(1-p_0)s < r(1-p_0)s$, so $T^*_G < T^*_\delta$ because F_G is decreasing. The other cases are analogous.

Step 5. For $T > 0$, $f_T(\lambda) := (r+\lambda)e^{-\lambda T}$ satisfies $f_T''(\lambda) = Te^{-\lambda T}\big[T(r+\lambda) - 2\big]$. If $T \geq 2/r$, then f_T is strictly convex on $(0,\infty)$, so $R(T) > 1$ by Jensen's inequality, and therefore $\hat{T} < 2/r$. If G is supported in $(0, \lambda_{\max}]$ and $T \leq 2/(r+\lambda_{\max})$, then f_T is strictly concave on $(0, \lambda_{\max}]$, so $R(T) < 1$, and therefore $\hat{T} > 2/(r+\lambda_{\max})$.

Step 6. $T_\delta^* = \hat{\lambda}^{-1} \ln\big[(g-s)p_0(r+\hat{\lambda})/(rs(1-p_0))\big]$ is continuous and strictly increasing in p_0, from 0 at $p_0 = \alpha(\hat{\lambda})$ to ∞ as $p_0 \to 1$. Solving $T_\delta^* = \hat{T}$ for p_0 gives $\hat{p}_0$. □

Proof of Proposition 2

Write L_i, ψ_i, χ_i, T_i^*, and V_i for the objects associated with G_i. Fix $T > 0$ and $\varepsilon > 0$. Using $1 - e^{-x} \leq \min\{x, 1\}$ and $xe^{-xT} \leq 1/(eT)$,

$$0 \leq 1 - L_i(T) \leq \varepsilon T + G_i\big((\varepsilon,\infty)\big),$$
$$0 \leq \int \lambda e^{-\lambda T} G_i(\mathrm{d}\lambda) \leq \varepsilon + \frac{G_i\big((\varepsilon,\infty)\big)}{eT}.$$

Letting $i \to \infty$ and then $\varepsilon \to 0$ yields $L_i(T) \to 1$ and $\int \lambda e^{-\lambda T} G_i(\mathrm{d}\lambda) \to 0$. Hence $\psi_i(T) \to r(gp_0 - s)$ for every $T > 0$.

(a) If $gp_0 > s$, then for every $\bar{T} > 0$, $\psi_i(\bar{T}) > 0$ for all large i, so $T_i^* > \bar{T}$ because ψ_i is decreasing. If $gp_0 < s$, then $\psi_i(\bar{T}) < 0$ for all large i, so $T_i^* < \bar{T}$.

(b) By Theorem 1(d), $V_i = p_0 g + e^{-rT_i^*}\chi_i(T_i^*) = \max_{T\geq 0}\{p_0 g + e^{-rT}\chi_i(T)\}$, with $|\chi_i| \leq g$, and $V_i \geq \max\{gp_0, s\}$ because the DM can use one arm forever. If $gp_0 > s$, then $T_i^* \to \infty$ and $V_i \to p_0 g$. If $gp_0 < s$, then $s \leq V_i = s + \int_0^{T_i^*} e^{-rT}\psi_i(T)\,\mathrm{d}T \leq s + T_i^*\,\psi_i(0)^+$, where $\psi_i(0) = (g-s)p_0(r+\hat{\lambda}) - r(1-p_0)s$ does not depend on i; hence $V_i \to s$. If $gp_0 = s$, then $\chi_i(T) = (1-p_0)s\,(1 - L_i(T))$. For any $\bar{T} > 0$, $e^{-rT}\chi_i(T) \leq (1-p_0)s\max\{1 - L_i(\bar{T}), e^{-r\bar{T}}\}$ for all T, whose limit superior is at most $(1-p_0)se^{-r\bar{T}}$; letting $\bar{T} \to \infty$ gives $V_i \to p_0 g = s$. □

Proof of Lemma 3

First, the clock is well defined. On each interval on which K_{-n} is Lipschitz, the Carathéodory equation $\mathrm{d}T_t/\mathrm{d}t = k_n(t) + K_{-n}(T_t)$ has a unique solution until T_t reaches the end of the interval. Since $\mathrm{d}T_t/\mathrm{d}t \leq N$, only finitely many intervals are crossed in any bounded time.

At the posterior μ_T, the problem is the original one with prior μ_T. The corresponding functions are $D(T+\cdot)/D(T)$, $\chi(T+\cdot)/D(T)$, and $\psi(T+\cdot)/D(T)$; for example, $(1-$

$p_T)s - (g-s)p_T L(T+\tau)/L(T) = \chi(T+\tau)/D(T)$. As in the proof of Lemma 2, player n's expected flow payoff at time t is $s + A_T(T_t) - k_n(t)\chi(T_t)/D(T)$, where $A_T(T') \coloneqq (g-s)\big(D(T) - D(T')\big)/D(T)$. Since $\mathrm{d}A_T(T_t)/\mathrm{d}t = \chi'(T_t)K_t/D(T)$ with $K_t = k_n(t) + K_{-n}(T_t)$, integration by parts gives

$$\begin{aligned} Z_n(T) &= \int_0^\infty e^{-rt}\big[\chi'(T_t)K_t - r\chi(T_t)k_n(t)\big]\mathrm{d}t \\ &= \int_0^\infty e^{-rt}\big[\chi'(T_t)K_{-n}(T_t) + \psi(T_t)k_n(t)\big]\mathrm{d}t. \end{aligned}$$

Averaging over players, $\frac{1}{N}\sum_n \big[\chi' K_{-n} + \psi k_n\big] = \frac{1}{N}\big[(N-1)\chi' + \psi\big]K = \big[\chi' - (r/N)\chi\big]K = \psi_N K$. □

Proof of Proposition 3

By Lemma 3, the planner maximizes $\int_0^\infty e^{-rt}\psi_N(T_t)\,\mathrm{d}T_t$ subject to $\mathrm{d}T_t/\mathrm{d}t \in [0,N]$. The function $\psi_N = (g-s)p_0\big((r/N)L - L'\big) - (r/N)(1-p_0)s$ is strictly decreasing with limit $-(r/N)(1-p_0)s < 0$. The argument in the proof of Theorem 1(b), with $T_t \le Nt$ and hence $\tau(T) \ge T/N$, shows that the average surplus is at most $\int_0^{T_N^*} e^{-rT/N}\psi_N(T)\,\mathrm{d}T$, with equality if and only if all players experiment until T_N^* (almost everywhere). This integral equals $e^{-rT_N^*/N}\chi(T_N^*) - \chi(0)$, which gives V_N^P. In normalized form, $\psi_N(T) = D(T)\big[(g-s)\bar\lambda_T - (r/N)(s-gp_T)\big]$, which is positive if and only if $(g-s)N\bar\lambda > r(s-gp)$, that is, $p > \alpha(Nm)$; with finite support this is $\sum_j \pi_j/\alpha(N\lambda_j) > 1$. Finally, $\psi_N = \psi + r(1-1/N)\chi$. At T_1^*, $\psi = 0$ and $r\chi = \chi' > 0$, so $\psi_N(T_1^*) > 0$ and $T_N^* > T_1^*$. □

Proof of Theorem 3

I first state two auxiliary results. For a fixed player, let K_{-n} denote the others' aggregate admissible strategy.

Lemma 4 (Verification). *Let $Z : [0,\infty) \to \mathbb{R}$ be bounded and continuously differentiable, and suppose that for every $T \ge 0$,*

$$rZ(T) = K_{-n}(T)\big[\chi'(T) + Z'(T)\big] + \max_{k\in[0,1]} k\big[\psi(T) + Z'(T)\big]. \tag{14}$$

Then for every state T and every measurable deviation, player n's surplus (12) is at most $Z(T)$. Equality holds for any admissible Markov policy that attains the maximum in (14) at every state.

Proof. Let T_t be the clock under a deviation $k_n(t)$, with $T_0 = T$. By (14), $rZ \geq K_{-n}(\chi' + Z') + k_n(\psi + Z')$ at every state. Hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big[e^{-rt}Z(T_t)\big] &= e^{-rt}\big[Z'(T_t)\big(k_n(t) + K_{-n}(T_t)\big) - rZ(T_t)\big] \\ &\leq -e^{-rt}\big[\chi'(T_t)K_{-n}(T_t) + \psi(T_t)k_n(t)\big],\end{aligned}$$

with equality if $k_n(t)$ attains the maximum at T_t. Integrate from 0 to t and let $t \to \infty$, using the boundedness of Z. □

Lemma 5 (Regularity and best-response conditions)**.** *Let all players use an admissible strategy k, and let $Z(T)$ be the resulting surplus at state T. Let (a, b) be an open interval on which k is Lipschitz continuous and positive.*

(i) Z is continuously differentiable on (a, b), and $rZ = NkZ' + k\big[(N-1)\chi' + \psi\big]$ there. Equivalently, $\psi + Z' = \frac{r}{Nk}\big[Z - (N-1)k\chi\big]$.

(ii) If k is a best response from every state in (a, b), then for every $T \in (a, b)$: $k(T) > 0 \Rightarrow \psi(T) + Z'(T) \geq 0$, and $k(T) < 1 \Rightarrow \psi(T) + Z'(T) \leq 0$.

Proof. (i) For $T < T'$ in (a, b), the equilibrium clock moves from T to T' in calendar time $\theta(T, T') = \int_T^{T'} \mathrm{d}\sigma/(Nk(\sigma))$, which is finite and continuously differentiable in T. By Lemma 3, changing variables from t to the clock,

$$Z(T) = \int_T^{T'} e^{-r\theta(T,\sigma)} \frac{(N-1)\chi'(\sigma) + \psi(\sigma)}{N}\,\mathrm{d}\sigma + e^{-r\theta(T,T')}Z(T').$$

Differentiating with respect to T and using $\partial\theta(T, \sigma)/\partial T = -1/(Nk(T))$ gives $Z'(T) = -[(N-1)\chi'(T) + \psi(T)]/N + rZ(T)/(Nk(T))$. This is the stated equation, and its right side is continuous. The equivalent form follows from $\chi' - \psi = r\chi$.

(ii) Fix $T \in (a, b)$ and $c \in [0, 1]$, and let player n use c on $[0, h)$ and revert to k thereafter. By Lemma 3, the deviation yields

$$Z_c(h) = \int_0^h e^{-rt}\big[\chi'(T_t)(N-1)k(T_t) + \psi(T_t)c\big]\,\mathrm{d}t + e^{-rh}Z(T_h),$$

where $\mathrm{d}T_t/\mathrm{d}t = c + (N-1)k(T_t)$ and $T_0 = T$. Since k is continuous at T and Z is differentiable there,

$$\begin{aligned}Z_c(h) = Z(T) + h\Big\{&(N-1)k(T)\big[\chi'(T) + Z'(T)\big] \\ &+ c\big[\psi(T) + Z'(T)\big] - rZ(T)\Big\} + o(h).\end{aligned}$$

By (i), the braces vanish when $c = k(T)$, so $Z_c(h) - Z(T) = h\,(c - k(T))\big[\psi(T) + Z'(T)\big] + o(h)$. Best-response behavior at T requires $Z_c(h) \le Z(T)$ for all small $h > 0$; take $c = 0$ if $k(T) > 0$ and $c = 1$ if $k(T) < 1$. □

Construction Let $H(T) := \int_T^{T_1^*} \psi(\tau)\,\mathrm{d}\tau - (N-1)\chi(T)$. On $[0, T_1^*]$, $H' = -\psi - (N-1)\chi' < 0$, and $H(T_1^*) = -(N-1)\chi(T_1^*) < 0$, because $r\chi(T_1^*) = \chi'(T_1^*) > 0$. So $T^\dagger$ is well defined. Define

$$Z^\dagger(T) := \begin{cases} N\chi(T^\dagger)e^{-r(T^\dagger - T)/N} - \chi(T), & 0 \le T < T^\dagger, \\ \int_T^{T_1^*} \psi(\tau)\,\mathrm{d}\tau, & T^\dagger \le T < T_1^*, \\ 0, & T \ge T_1^*. \end{cases}$$

I claim that $Z^\dagger$ is continuously differentiable on $[0, \infty)$ and that on $[0, T_1^*)$ it satisfies

$$Z'(T) = -\psi(T) + \frac{r}{N}\big[Z(T) - (N-1)\chi(T)\big]^+, \tag{15}$$

with $Z^\dagger < (N-1)\chi$ on $(T^\dagger, T_1^*)$ and $Z^\dagger > (N-1)\chi$ on $[0, T^\dagger)$. On $[T^\dagger, T_1^*)$, $Z^\dagger - (N-1)\chi = H \le 0$ and $(Z^\dagger)' = -\psi$. The inequality is strict for $T > T^\dagger$. At $T^\dagger$ it is an equality when $T^\dagger > 0$, and also when $T^\dagger = 0$ and $H(0) = 0$; it is strict when $T^\dagger = 0$ and $H(0) < 0$. On $[0, T^\dagger)$,

$$(Z^\dagger)' = \tfrac{r}{N}\big(Z^\dagger + \chi\big) - \chi' = -\psi + \tfrac{r}{N}\big(Z^\dagger - (N-1)\chi\big),$$

and $Z^\dagger - (N-1)\chi = N\big[\chi(T^\dagger)e^{-r(T^\dagger - T)/N} - \chi(T)\big] > 0$, because $T \mapsto e^{-rT/N}\chi(T)$ has derivative $e^{-rT/N}\psi_N(T) > 0$ on $[0, T_1^*) \subset [0, T_N^*)$. If $T^\dagger > 0$, both formulas give $Z^\dagger = (N-1)\chi(T^\dagger)$ and $(Z^\dagger)' = -\psi(T^\dagger)$ at the joining point. If $T^\dagger = 0$, only the integral formula is used at the left endpoint, and its right derivative is $-\psi(0)$. At T_1^*, $Z^\dagger \to 0$ and $(Z^\dagger)' \to -\psi(T_1^*) = 0$ from the left. Note that $k^\dagger = \min\{1, Z^\dagger/((N-1)\chi)\}$ on $[0, T_1^*)$ wherever $\chi > 0$, and $k^\dagger = 1$ wherever $\chi \le 0$, since $\chi \le 0$ only on $[0, T^\dagger)$.

Existence Let the other players use $k^\dagger$, so $K_{-n} = (N-1)k^\dagger$. I verify (14) for $Z^\dagger$, with the maximum attained at $k^\dagger$. For $T \ge T_1^*$: $Z^\dagger = (Z^\dagger)' = 0$, $k^\dagger = 0$, and $\max_k k\psi(T) = 0$ because $\psi(T) \le 0$. On the mixing region, $\psi + (Z^\dagger)' = 0$, so every k attains the maximum, and the right side of (14) equals $(N-1)k^\dagger(\chi' - \psi) = (N-1)k^\dagger r\chi = rZ^\dagger$. On $[0, T^\dagger)$, $\psi + (Z^\dagger)' = \frac{r}{N}\big(Z^\dagger - (N-1)\chi\big) > 0$, so $k = 1$ attains the maximum, and the right side equals $(N-1)\chi' + \psi + N(Z^\dagger)' = (N-1)(\chi' - \psi) + rZ^\dagger - (N-1)r\chi = rZ^\dagger$. By Lemma 4, $k^\dagger$ is a best response from every state, and $Z^\dagger$ is the equilibrium surplus. The strategy $k^\dagger$ is continuous and Lipschitz: it is constant on each nonmixing interval; on the mixing

interval, its numerator and denominator are Lipschitz because ψ and χ' are bounded, and $\chi \geq \chi(T^\dagger) > 0$. Thus it is admissible.

Properties (a) On $[T^\dagger, T_1^*)$, the numerator of $k^\dagger$ is positive and strictly decreasing, and the denominator is positive and increasing. Players mix only where $\chi > 0$, that is, $p_T < s/g$. (b) By the Lipschitz property established above and $k^\dagger(T_1^*) = 0$, T_1^* is a rest point of the equation $\mathrm{d}T_t/\mathrm{d}t = Nk^\dagger(T_t)$ on $[T^\dagger, \infty)$, and by uniqueness of solutions a path started below T_1^* never reaches it. Since T_t is increasing and bounded by T_1^*, it converges to a rest point of $k^\dagger$, and $k^\dagger > 0$ on $[0, T_1^*)$, so the limit is T_1^*. (c) $V_N^E = s + Z^\dagger(0)$ because $D(0) = 1$. If $T^\dagger = 0$, then $Z^\dagger(0) = \int_0^{T_1^*} \psi > \int_0^{T_1^*} e^{-rT}\psi = V_1 - s$. If $T^\dagger > 0$, consider the deviation "experiment until the clock reaches T_1^*, then stop" (see step (i) below). It yields more than $V_1 - s$, because the others' experimentation on $[0, T^\dagger)$ makes the clock reach T_1^* strictly before calendar time T_1^*. Hence $V_N^E > V_1$. Finally, by Lemma 3, the equilibrium average surplus is $\int e^{-rt}\psi_N(T_t)\,\mathrm{d}T_t$ along a path that differs from the efficient one, so $V_N^E < V_N^P$ by the uniqueness part of Proposition 3.

Uniqueness Let k be the common strategy of a symmetric MPE and Z the resulting surplus.

(i) $Z \geq 0$, *and* $Z(T) \geq e^{-r(T_1^* - T)}\chi(T_1^*) - \chi(T) > 0$ *for* $T < T_1^*$. The deviation $k_n \equiv 0$ yields $\int e^{-rt}\chi' K_{-n}\,\mathrm{d}t \geq 0$. Consider instead the deviation of experimenting fully until the clock reaches T_1^*, at calendar time $\tau_1 \leq T_1^* - T$, and never afterwards. On $[0, \tau_1)$, $\chi' K_{-n} + \psi = \chi'(1 + K_{-n}) - r\chi$, so $e^{-rt}[\chi' K_{-n} + \psi] = \frac{\mathrm{d}}{\mathrm{d}t}\left[e^{-rt}\chi(T_t)\right]$. The deviation therefore yields at least $e^{-r\tau_1}\chi(T_1^*) - \chi(T) \geq e^{-r(T_1^* - T)}\chi(T_1^*) - \chi(T)$, using $\chi(T_1^*) > 0$. This lower bound is the single-agent surplus $\int_T^{T_1^*} e^{-r(\tau - T)}\psi(\tau)\,\mathrm{d}\tau > 0$.

(ii) $k > 0$ *on* $[0, T_1^*)$, *and the clock crosses every point of* $[0, T_1^*)$ *in finite time.* If $k(T_0) = 0$, then T_0 is a rest point: k is Lipschitz on some $[T_0, T_0 + \varepsilon)$ with $k(T_0) = 0$, so the clock started at T_0 stays there. Hence $Z(T_0) = 0$, contradicting (i). Next, let $b < T_1^*$ be a breakpoint of k, and suppose that $k(T) \to 0$ as $T \uparrow b$. Then the clock started below b approaches b only asymptotically, and $|Z(T)| \leq C \sup_{[T,b)} k \to 0$ as $T \uparrow b$, where $C := \sup_T \left[(N-1)|\chi'(T)| + |\psi(T)|\right]/r < \infty$. This contradicts the lower bound in (i), which converges to the positive number $e^{-r(T_1^* - b)}\chi(T_1^*) - \chi(b)$. Hence the left limit $k(b-)$ is positive at every breakpoint $b < T_1^*$, so the clock crosses each such breakpoint in finite time, and Z is continuous on $[0, T_1^*)$.

(iii) $k = 0$ *on* $[T_1^*, \infty)$. Suppose $k(T_0) > 0$ for some $T_0 \geq T_1^*$, and let $T_b := \inf\{T > T_0 : k(T) = 0 \text{ or } k(T-) = 0\} \in (T_0, \infty]$. By admissibility, $[T_0, T_b)$ contains finitely many

breakpoints of k on every compact set; k and its left limits are positive there, so the clock crosses each breakpoint in finite time and Z is continuous on $[T_0, T_b)$. By Lemma 5, $Z' \geq -\psi > 0$ at all other points of (T_1^*, T_b), so Z is strictly increasing on $[T_0, T_b)$. If $T_b = \infty$, this contradicts the boundedness of Z, since $-\psi(T) \geq -\psi(T_0 + 1) > 0$ for $T \geq T_0 + 1$. If $T_b < \infty$, then $Z(T) \to 0$ as $T \uparrow T_b$. Either $k(T_b-) > 0$ and $k(T_b) = 0$, in which case the clock reaches the rest point T_b in a time that vanishes as $T \uparrow T_b$ and $Z(T_b) = 0$; or $k(T_b-) = 0$, in which case $|Z(T)| \leq C \sup_{[T,T_b)} k \to 0$. Since Z is strictly increasing on $[T_0, T_b)$, $Z(T_0) < 0$, contradicting (i).

(iv) *The equilibrium surplus solves* (15) *on* $[0, T_1^*)$. By (ii), Lemma 5 applies on the interior of every Lipschitz piece of k in $[0, T_1^*)$. There, the best-response conditions and the identity $\psi + Z' = \frac{r}{Nk}[Z - (N-1)k\chi]$ give the following. If $0 < k < 1$, then $Z = (N-1)k\chi$, which requires $\chi > 0$. If $k = 1$, then $Z \geq (N-1)\chi$. So $k = \min\{1, Z/((N-1)\chi)\}$ where $\chi > 0$, and $k = 1$ where $\chi \leq 0$. Substituting into the identity yields (15) at all but finitely many points of every compact subset of $[0, T_1^*)$. Since Z is continuous and the right side of (15) is continuous along Z, Z is continuously differentiable on $[0, T_1^*)$ and satisfies (15) everywhere there.

(v) $Z(T) \to 0$ *as* $T \uparrow T_1^*$. By (ii) and (iii), the clock started below T_1^* increases toward the rest point T_1^*. Either it reaches T_1^* in a time that vanishes as $T \uparrow T_1^*$, if $\lim_{T \uparrow T_1^*} k(T) > 0$, or $k(T) \to 0$; in both cases $Z(T) \to 0$.

(vi) *Conclusion.* The right side of (15), $F(T, z) := -\psi(T) + \frac{r}{N}[z - (N-1)\chi(T)]^+$, is continuous on $[0, T_1^*] \times \mathbb{R}$ and Lipschitz in z with constant r/N. By (iv) and (v), Z extends continuously to $[0, T_1^*]$ and solves $Z(T) = -\int_T^{T_1^*} F(\tau, Z(\tau))\, d\tau$. By the Picard–Lindelöf theorem, this terminal-value problem has a unique solution, which by the construction is $Z^\dagger$. Hence $Z = Z^\dagger$ on $[0, T_1^*]$, and by (iv), $k = k^\dagger$ on $[0, T_1^*)$. Together with (iii), $k = k^\dagger$.

The case $p_0 \leq \alpha(\hat{\lambda})$ Then $T_1^* = 0$. The strategy $k \equiv 0$ with $Z \equiv 0$ satisfies (14), because $\psi \leq 0$. Step (iii) applies verbatim and shows that every symmetric MPE has $k \equiv 0$. □

Proof of Corollary 4

$T_1^* < T_N^*$ by Proposition 3. On $(T^\dagger, T_1^*)$, $Nk^\dagger < N$, and $Nk^\dagger(T) \to 0$ as $T \uparrow T_1^*$. □

Proof of Proposition 4

(a) By Theorem 3(b), the equilibrium clock converges to T_1^* without reaching it. The probability that a good arm is never developed is therefore $\lim_t L(T_t) = L(T_1^*)$, and beliefs converge to $\mu_{T_1^*}$, where $p_{T_1^*} = \alpha(m_{T_1^*})$ by the argument of Proposition 1(c). These are exactly the objects compared in Proposition 1(c)–(d) and Theorem 2.

(b) From $\psi(T_1^*) = 0$, $(g-s)p_0 L(T_1^*)(r + m_{T_1^*}) = r(1-p_0)s$. From $\psi_N(T_N^*) = 0$, $(g-s)p_0 L(T_N^*)(r + Nm_{T_N^*}) = r(1-p_0)s$. Dividing gives the stated ratio. Since $0 < T_1^* < T_N^*$ and m is strictly decreasing with $m_0 = \hat{\lambda}$, we have $m_{T_N^*} < m_{T_1^*} < \hat{\lambda}$. Since $x \mapsto (r+Nx)/(r+x)$ is strictly increasing,

$$\frac{r + Nm_{T_N^*}}{r + m_{T_1^*}} < \frac{r + Nm_{T_1^*}}{r + m_{T_1^*}} < \frac{r + N\hat{\lambda}}{r + \hat{\lambda}}.$$

(c) For every p_0, $\max\{gp_0, s\} \leq V_1 \leq V_N^E \leq V_N^P$. The first inequality holds because the DM can always use one arm forever, the second by step (i) of the uniqueness proof, and the third by the definition of the planner. By Proposition 3, $V_N^P = \max_T\{p_0 g + e^{-(r/N)T}\chi(T)\}$ is the value of a single DM with discount rate r/N. Proposition 2(b), whose proof holds for any discount rate, yields $V_N^P \to \max\{gp_0, s\}$ along (G_i). □

Proof of Proposition 5

At time t, player n's expected flow payoff is $g(\lambda)$ on the event that the arm is good with rate λ and a breakthrough has occurred, which has ex ante probability $p_0(1 - e^{-\lambda T_t})G(\mathrm{d}\lambda)$. On the event of silence, which has probability $D(T_t)$, it is $(1-k_n)s + k_n p_0 \int g(\lambda)e^{-\lambda T_t}G(\mathrm{d}\lambda)/D(T_t)$. Subtracting s, the expected flow surplus is

$$\begin{aligned} &p_0 \int \big(g(\lambda) - s\big)\big(1 - e^{-\lambda T_t}\big)G(\mathrm{d}\lambda) \\ &\quad - k_n\Big[(1-p_0)s - p_0 \int \big(g(\lambda) - s\big)e^{-\lambda T_t}G(\mathrm{d}\lambda)\Big]. \end{aligned}$$

With $\tilde{g} = s + c$ and $\tilde{L}(T) = \int e^{-\lambda T}\tilde{G}(\mathrm{d}\lambda)$, this equals $(\tilde{g}-s)p_0\big(1 - \tilde{L}(T_t)\big) - k_n\tilde{\chi}(T_t)$, where $\tilde{\chi}(T) := (1-p_0)s - (\tilde{g}-s)p_0\tilde{L}(T)$. This is precisely the expected flow surplus in the model of Assumption 1 with $(\tilde{g}, \tilde{G})$, for every profile. The equivalence follows. With finitely many arrival rates, the index condition $\tilde{\psi} > 0$, divided by $D(T)$ and written in terms of posterior probabilities π_j, reads $\sum_j \pi_j(r + \lambda_j)(g(\lambda_j) - s) > rs\big(1 - \sum_j \pi_j\big)$, that is, $\sum_j \pi_j/\alpha_j > 1$. □

One-shot prizes

The payoff of a policy is $\int_0^\infty e^{-rt} D(T_t) \big[\bar{\lambda}_{T_t} W - \kappa\big] \mathrm{d}T_t$, whose integrand need not be monotone in the clock but changes sign at most once, from positive to negative, because $\bar{\lambda}_T$ falls with silence (Lemma 1(c)); the argument of Theorem 1 needs only this single crossing.

B Numerical methods

All figures are computed from the explicit characterizations in the text. Roots of ψ, ψ_N, and H are found by bracketing and Brent's method; integrals of ψ and χ are evaluated analytically, using the Laplace transform of the two-point or gamma distribution. Equilibrium paths in calendar time are obtained by integrating $\mathrm{d}T_t/\mathrm{d}t = Nk^\dagger(T_t)$ with a second-order Runge–Kutta scheme. Three independent checks, which do not use these explicit characterizations to solve the decision problems, support the results.

(i) *Payoff representation.* Monte Carlo simulation of actual lump-sum payoffs (2×10^5 draws) reproduces the payoff of experimenting until a fixed clock, $W(\tau) = p_0 g + e^{-r\tau}\chi(\tau)$, within two standard errors.

(ii) *Single-agent boundary.* Discrete-time dynamic programming ($\Delta t = 0.01$) on a 161×161 grid of the two-rate belief triangle, with $\lambda^- = 1$, $\lambda^+ = 10$, $g = 2$, $s = 1$, $r = 1$, recovers the stopping boundary ℓ_1 to within 1.5 grid steps, and values within 0.002 of (8).

(iii) *Equilibrium.* Given that the other players use $k^\dagger$, a player's best-response problem along the no-news path was solved by discrete-time dynamic programming ($\Delta t = 5 \times 10^{-4}$). The cases checked are two-rate and gamma priors and the known-rate case, with $N \in \{2, 3, 4\}$. The best-response value matches $s + Z^\dagger/D$ to within 10^{-4}. The gain from experimenting rather than playing safe is strictly positive where $k^\dagger = 1$, strictly negative where $k^\dagger = 0$, and zero up to 2×10^{-7} in the mixing region, as indifference requires. Path integrals of (12) along the equilibrium path reproduce V_N^E to within 2×10^{-6}.

The supplementary computational material gives instructions for running the replication archive. Its package versions and commands document reproducible calculations with Python, NumPy, SciPy, and Matplotlib. The figures visualize model calculations, not

observed data or generated raster illustrations. Opus 5.5 (Anthropic) assisted with figure design and drafting the Python figure-generation code.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interests

The author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and code availability

This theoretical study uses no externally collected data. The accompanying replication archive contains the Python model implementation, figure-generation scripts, numerical verification scripts, and computational outputs used for the figures and numerical checks.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, I used Opus 5.5 (Anthropic) for assistance with drafting the abstract, introduction, and conclusion; Example 1 ("Gamma-distributed arrival rates"); selected lemma, proposition, theorem, and section titles; and the one-shot-prize discussion in Section 6.3. Opus 5.5 also assisted with figure design and drafting the associated Python code, and with a revision of the exposition throughout the manuscript (including the description of the problem and the motivating examples) that did not change the mathematical content of the formal results or proofs. GPT-6 Astra (OpenAI) provided substantial assistance with the formalization and proof of Proposition 4 ("Free riding and rate–size uncertainty"). OpenAI Codex assisted with preparing the submission package, including LaTeX formatting, bibliography styling, disclosure drafting, file packaging, and automated document-consistency and replication

checks. After using these tools, I reviewed and edited the content as needed, and I take full responsibility for the final manuscript.

References

Anat R. Admati and Motty Perry. Joint projects without commitment. *Review of Economic Studies*, 58(2):259–276, 1991. doi: 10.2307/2297967.

Dirk Bergemann and Juuso Välimäki. Bandit problems. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics*, volume 1, pages 336–340. Palgrave Macmillan, Basingstoke, second edition, 2008. doi: 10.1057/9780230226203.0088.

Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2): 349–374, 1999. doi: 10.1111/1468-0262.00022.

Svetlana Boyarchenko. Super- and submodularity of stopping games with random observations. *Economic Theory*, 70(4):983–1022, 2020. doi: 10.1007/s00199-019-01198-4.

Arthur Campbell, Florian Ederer, and Johannes Spinnewijn. Delay and deadlines: Freeriding and information revelation in partnerships. *American Economic Journal: Microeconomics*, 6(2):163–204, 2014. doi: 10.1257/mic.6.2.163.

Andrew Caplin and John Leahy. Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of Economics*, 116(1):55–79, 2001. doi: 10.1162/003355301556347.

Jay Pil Choi. Dynamic R&D competition under "hazard rate" uncertainty. *RAND Journal of Economics*, 22(4):596–610, 1991. doi: 10.2307/2600992.

Y. S. Chow, Herbert Robbins, and David Siegmund. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, Boston, 1971.

Asaf Cohen and Eilon Solan. Bandit problems with Lévy processes. *Mathematics of Operations Research*, 38(1):92–107, 2013. doi: 10.1287/moor.1120.0564.

Partha Dasgupta and Eric Maskin. Uncertainty and hyperbolic discounting. *American Economic Review*, 95(4):1290–1299, 2005. doi: 10.1257/0002828054825637.

Nicole El Karoui and Ioannis Karatzas. Dynamic allocation problems in continuous time. *The Annals of Applied Probability*, 4(2):255–286, 1994. doi: 10.1214/aoap/1177005062.

Jeffrey C. Ely and Martin Szydlowski. Moving the goalposts. *Journal of Political Economy*, 128(2):468–506, 2020. doi: 10.1086/704387.

Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12):3651–3684, 2018. doi: 10.1257/aer.20150742.

George Georgiadis. Projects and team dynamics. *Review of Economic Studies*, 82(1):187–218, 2015. doi: 10.1093/restud/rdu031.

John Gittins, Kevin Glazebrook, and Richard Weber. *Multi-armed Bandit Allocation Indices*. Wiley, Chichester, second edition, 2011. doi: 10.1002/9780470980033.

John C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):148–177, 1979. doi: 10.1111/j.2517-6161.1979.tb01068.x.

Francisco M. Gonzalez and Shouyong Shi. An equilibrium theory of learning, search, and wages. *Econometrica*, 78(2):509–537, 2010. doi: 10.3982/ECTA8061.

James Heckman and Burton Singer. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320, 1984. doi: 10.2307/1911491.

Josef Hofbauer and Karl Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, 1998. doi: 10.1017/CBO9781139173179.

Johannes Hörner and Andrzej Skrzypacz. Learning, experimentation, and information design. In Bo Honoré, Ariel Pakes, Monika Piazzesi, and Larry Samuelson, editors, *Advances in Economics and Econometrics: Eleventh World Congress*, volume 1, pages 63–98. Cambridge University Press, Cambridge, 2017. doi: 10.1017/9781108227162.002.

Haya Kaspi and Avishai Mandelbaum. Multi-armed bandits in discrete and continuous time. *The Annals of Applied Probability*, 8(4):1270–1290, 1998. doi: 10.1214/aoap/1028903380.

Godfrey Keller and Sven Rady. Strategic experimentation with Poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010. doi: 10.3982/TE595.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005. doi: 10.1111/j.1468-0262.2005.00564.x.

Philipp Külpmann. Projects with uncertain requirements and deadlines. *Economic Theory*, 81:835–864, 2026. doi: 10.1007/s00199-025-01668-y.

Tony Lancaster. *The Econometric Analysis of Transition Data*. Number 17 in Econometric Society Monographs. Cambridge University Press, Cambridge, 1990. doi: 10.1017/CCOL0521265967.

David A. Malueg and Shunichi O. Tsutsui. Dynamic R&D competition with learning. *RAND Journal of Economics*, 28(4):751–772, 1997. doi: 10.2307/2555785.

Leslie M. Marx and Steven A. Matthews. Dynamic voluntary contribution to a public project. *Review of Economic Studies*, 67(2):327–358, 2000. doi: 10.1111/1467-937X.00134.

Tatiana Mayskaya. Following beliefs or excluding the worst? The role of unfindable state in learning. *European Economic Review*, 162:104653, 2024. doi: 10.1016/j.euroecorev.2023.104653. Earlier version circulated as "Dynamic Choice of Information Sources," Working Paper WP9/2019/05, National Research University Higher School of Economics, Moscow, 2019.

Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*. Number 293 in Grundlehren der mathematischen Wissenschaften. Springer, Berlin, third edition, 1999. doi: 10.1007/978-3-662-06400-9.

Moshe Shaked and J. George Shanthikumar. *Stochastic Orders*. Springer Series in Statistics. Springer, New York, 2007. doi: 10.1007/978-0-387-34675-5.

Peter D. Sozou. On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1409):2015–2020, 1998. doi: 10.1098/rspb.1998.0534.

Martin L. Weitzman. Why the far-distant future should be discounted at its lowest possible rate. *Journal of Environmental Economics and Management*, 36(3):201–208, 1998. doi: 10.1006/jeem.1998.1052.