Strategic Experimentation with Evolving Resource Quality

Faheem Gilani* Caroline Thomas

Working Draft: October 2026

Abstract

We study strategic experimentation on a resource that, if productive, depletes when harvested and grows deterministically otherwise. Harvest opportunities arrive according to a poisson process whose rate is proportional to total experimentation intensity and whose value depends on the level of the resource at the time of arrival. In a symmetric Markov perfect equilibrium, when players are sufficiently confident that the resource is productive they exploit it at levels where a social planner would have them wait. When they are unsure, they stop experimenting at a higher belief than the planner would, whatever the initial resource level. In between, they exploit the resource to the point where it pays the same as an outside safe option, dissipating its rent as in an open-access fishery. Sustaining this requires Filippov solutions.

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

1 Introduction

A fishing guide who tries an unfamiliar site learns from her encounters whether the target species is there, but each encounter also disturbs the fish and makes them harder to catch for a time. In a controlled experiment with rainbow trout, Koeck et al. (2020) find that temporary closures raise later catch rates, so catchability recovers when a site is left alone.[1] Phishing works in a similar way. Whether or not a victim who takes the bait ends up paying, the reports and warnings that follow can tell other scammers that the population is vulnerable, but they also make the population harder to fool until awareness fades.

In both examples, each encounter with the resource results in both a positive externality (public information) and a negative externality (the resource becomes harder to exploit for a time). The first effect pushes players to free ride on each other's experiments, as in the literature on strategic experimentation (Bolton and Harris, 1999; Keller et al., 2005). The second leads them to rush, as in the tragedy of the commons (Gordon, 1954; Levhari and Mirman, 1980). The two distortions push in opposite directions, and we ask which one dominates, where it dominates, and what equilibrium experimentation looks like when both are present.

We study a game of strategic experimentation with exponential bandits whose reward follows a piecewise deterministic controlled process. There are $N \geq 2$ players experimenting with a resource that is either productive or unproductive. If the resource is productive, then reward *opportunities* arrive according to a Poisson process with rate proportion to the total intensity of experimentation. Each oppotunity translates into a fixed *payoff* with probability equal to the current *quality* of the resource. Quality grows deterministically in between arrivals and resets to zero following any opportunity arrival, whether or not the opportunity translates into a payoff. At each moment in time, players divide a unit of effort between experimentation or exploitation of the resource and a safe activity that yields deterministic flow payoff. Actions and opportunities are publi (an extension where only payoffs are public is discussed in Section 8). Player strategies depend on the state of the game, which consists of the players' common posterior belief about the resource's productivity potential and its quality level. We solve the planner's problem from every state and, under conditions that hold in large groups construct a symmetric Markov perfect equilibrium on the whole state space.

Which distortion dominates depends above all on what the players believe (Theorem 3). When players are confident that the resource is productive, they use full effort at qualities at which a planner would let it recover. There is little left to learn, so the tragedy of the commons dominates. When players are unsure and quality is high, the comparison reverses. Players then experiment less than the planner, and their learning stops at a higher belief. A discovery is worth less to a player than to the planner for two reasons. First, a player values it only for herself. Second, after a discovery the players compete for the known resource and harvest it too early, which dissipates part of its rent. Figure 1 shows where each distortion applies. If failed opportunities are hidden,

[1] The experiment speaks only to recovery. The model assumes that encounters are public, that they occur at a rate that does not depend on catchability, and that each one fully resets catchability.

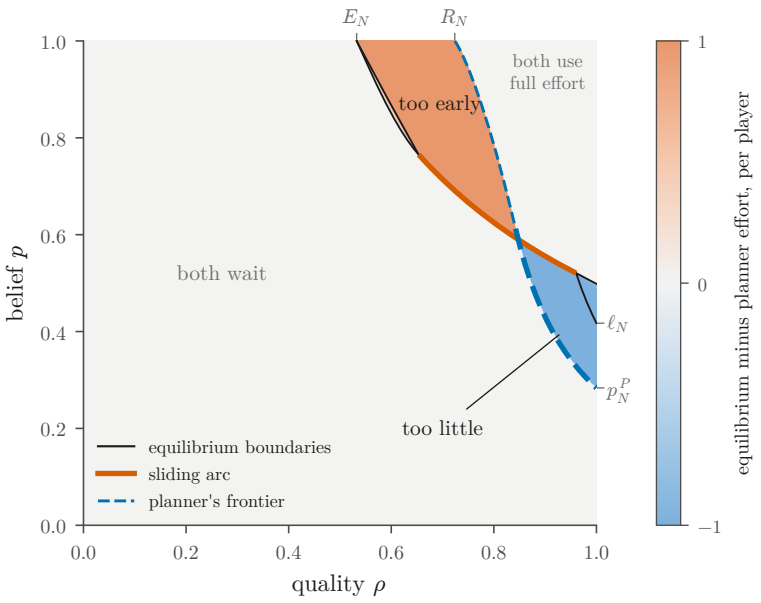


Figure 1: Players exploit the resource too early when they are confident that it is productive, and learn too little when they are unsure and quality is high. Each point is a state, with quality ρ on the horizontal axis and the players' belief p that the resource is productive on the vertical axis. Shading shows equilibrium minus efficient aggregate effort, per player. It is red where players use more effort than the planner (too early), blue where they use less (too little), and grey where both choose the same effort. Black curves bound the equilibrium's waiting and full-effort regions, the vermilion curve is the sliding arc, on which working the resource pays exactly the safe return, and the dashed blue curve bounds the planner's waiting region. On the top edge, E_N and R_N are the qualities at which the equilibrium and the planner start exploiting a resource known to be productive. On the right edge, ℓ_N and p_N^P are the beliefs at which equilibrium and efficient learning stop at full quality. Baseline parameters: $N = 3$ players, $\lambda = 1$, $h = 2$, $s = 1$, $\eta = 1$ and $r = 0.2$.

competition can dissipate all of the rent, and learning then stops at an even higher belief (Section 8).

How much is learned does not depend on the initial quality of the resource, either in equilibrium or in the planner's solution (Theorems 1 and 3). An idle resource recovers, so experimentation can always wait for better quality, and the decision to stop is effectively taken at full quality. Total experimentation and the probability of a discovery are then pinned down by the belief at which learning stops, and initial quality affects only when the experimenting takes place. Learning stops at the same belief in every Nash equilibrium in which players commit to deterministic effort schedules until the first opportunity and follow the known-state equilibrium afterward (Theorem 6), so this conclusion is not specific to the equilibrium we construct.

In the equilibrium we construct (Theorem 2 and Figure 5), nobody experiments while quality is low. Players use interior effort only in two narrow bands, one just above the quality at which a known productive resource starts to be exploited and one near full quality, where the equilibrium corresponds to the symmetric equilibrium of Keller et al. (2005). At intermediate qualities, players wait whenever working the resource would pay less than the safe activity and use full effort whenever

it would pay more. The state is therefore held on the curve along which working the resource pays exactly the safe return, and it slides down this curve as quality recovers. On the curve, recovery pins down aggregate effort, which does not depend on the number of players, and the flow rent of the resource is dissipated, as in the open-access fishery of Gordon (1954). Values stay above the safe payoff, because a discovery is still valuable. Every regular symmetric equilibrium behaves this way at these qualities (Proposition 5). Regularity requires a continuously differentiable value and a strategy that switches on finitely many smooth curves, waits at low beliefs and uses full effort at high ones. As the group grows, the range of such qualities expands to every quality at which a productive resource pays more than the safe activity (Proposition 4).

In this model, some deviations to markov strategies admit no ordinary trajectory best reponse (Proposition 5). Following Jaakkola and Wagener (2026), we use trajectories in the sense of Filippov (1988), which let the opponents' aggregate effort on the curve take any value between its limits on the two sides, in particular the value that keeps the state on it. If players revise their effort only at dates Δ apart, every deviation has an ordinary trajectory. As $\Delta \to 0$, the largest gain from a deviation from the constructed strategy vanishes, uniformly over compact sets of initial states with uncertain beliefs and quality below one, so the strategy yields approximate Nash equilibria, and the path of play before a discovery converges to the equilibrium path, which slides along the curve (Proposition 6). Whether an exact Markov perfect equilibrium with ordinary trajectories exists is left open.

A discovery ends all uncertainty about productivity. As a result, a player's payoff above the safe payoff, divided by the probability that the resource is unproductive, depends only on the deterministic path that the belief and quality follow without a discovery (Lemma 6.1). The game before a discovery thus becomes a deterministic differential game in the log odds of the belief and a clock that measures quality. In this game, each player's private gain from effort satisfies a linear transport equation along paths of waiting and of full effort (Lemma 6.2), which lets us build the equilibrium region by region and confirm it with a verification theorem for Filippov trajectories (Theorem 4). For the planner with an uncertain prior, the objective written in terms of the probability of no arrival conditional on productivity is strictly concave, which identifies the optimal policy (Theorem 1).

Related literature. Strategic experimentation with public information goes back to Bolton and Harris (1999), with Brownian news, and Keller et al. (2005), whose exponential-bandit game supplies our information structure. At full quality our equilibrium corresponds to their symmetric equilibrium, with the value of the known resource as the continuation after a success. Keller and Rady (2010) study inconclusive Poisson news, and Hörner and Skrzypacz (2017) survey the literature. Later work studies negatively correlated arms (Klein and Rady, 2011), bad news (Keller and Rady, 2015), private payoffs (Heidhues et al., 2015), collaboration and disagreement (Bonatti and Hörner, 2011, 2017), contests (Halac et al., 2017), networks (Board and Meyer-ter Vehn, 2024), and ways to overcome free riding (Hörner et al., 2022). In all of these papers, the payoff to a success does not

depend on a stock that successes deplete.

The papers closest to ours are Cripps and Thomas (2019) and Thomas (2021), which combine learning with congestion. In the queueing model of Cripps and Thomas (2019), observing other agents is informative but waiting behind them is costly. In Thomas (2021), players with private risky arms compete for a common safe arm, and congestion gives occupying it a strategic option value. In our model the congested object is the risky resource itself, which experiments both reveal and deplete, and the competition over it continues after its productivity is known.

The tension between preemption and free riding also appears in timing games. In preemption games, competition to move first dissipates the advantage of moving first (Fudenberg and Tirole, 1985), and several papers combine preemption with the incentive to wait (Park and Smith, 2008; Anderson et al., 2017). Kim and Lee (2014) study information acquisition in a war of attrition. In Eilat et al. (2026), two players repeatedly inspect, at a cost, whether a prize has appeared. After a failed inspection, the probability that it has appeared restarts from zero, much as quality does after an opportunity in our model. Depending on how much a player gains when the other finds the prize, competition to be first or free riding dominates, and players inspect too often or too rarely. In our model both forces operate in the same equilibrium, and which one dominates depends above all on the players' belief.

The commons literature studies the exploitation of a stock of known productivity (Gordon, 1954; Levhari and Mirman, 1980; Dockner et al., 2000). Antoniadou et al. (2013) and Fesselmeyer and Santugini (2013) add uncertainty about the stock or its environment, but their players do not learn by experimenting. On the methodological side, we use the Markov perfect equilibrium of Maskin and Tirole (2001) with Filippov trajectories, as Jaakkola and Wagener (2026) do for deterministic differential games with discontinuous Markov strategies, and Cortés (2008) surveys solution concepts for discontinuous dynamical systems. Our game has Poisson jumps and time-varying dynamics between them, and we prove existence by constructing the equilibrium and verifying it directly.

2 Model

2.1 Environment

There are $N \geq 2$ players and a common discount rate $r > 0$. At each date $t \geq 0$, player i chooses risky effort $a_t^i \in [0,1]$ and devotes the remaining $1 - a_t^i$ to a safe activity with flow payoff $s > 0$. Aggregate effort is $A_t = \sum_i a_t^i \in [0, N]$.

Productivity and opportunities. A fixed state of the world $\theta \in \{0,1\}$ determines whether the resource is productive ($\theta = 1$), that is, whether it can generate opportunities at all. If $\theta = 1$, player i's effort generates *opportunities* at Poisson rate λa_t^i, with $\lambda > 0$. If $\theta = 0$, there are no opportunities. An opportunity pays its finder a prize $h > 0$ with probability $\rho_t \in [0,1]$, the current *quality* of the resource, and nothing otherwise. Every opportunity, whether or not it pays, resets

quality to zero. Between opportunities quality recovers deterministically,

$$\dot{\rho}_t = \eta(1 - \rho_t), \qquad \eta > 0. \tag{2.1}$$

Information. Actions and opportunities are public. Let p_t be the common posterior probability that $\theta = 1$. The first opportunity, which we call a *discovery*, reveals $\theta = 1$. In the absence of opportunities, Bayes' rule gives

$$\dot{p}_t = -\lambda A_t p_t (1 - p_t). \tag{2.2}$$

Every opportunity therefore sends the state (p, ρ) to $(1, 0)$. The state space is $[0, 1]^2$. Following a discovery, the state remains on the $\{p = 1\}$ edge. The $\{\rho = 1\}$ edge is not reached in finite time from $\rho_0 < 1$, though it is an allowed initial state.

Payoffs. Write $g = \lambda h$ for the expected prize rate of full effort at full quality, and assume $0 < s < g$, so that the risky activity can be worthwhile. Put

$$q = \frac{s}{g} \in (0, 1), \qquad \beta = \frac{r}{\eta}, \qquad K = N - 1.$$

Here q is the quality at which a resource known to be productive pays exactly the safe flow, β is the discount rate relative to the rate of recovery, and K is the number of a player's opponents. Conditional on the public history, player i's expected prize rate is $\lambda a_t^i p_t \rho_t h = g p_t \rho_t a_t^i$. Player i's normalized payoff is

$$J^i = \mathbb{E} \int_0^\infty r e^{-rt} \left[s + a_t^i \left(g p_t \rho_t - s \right) \right] \mathrm{d}t. \tag{2.3}$$

The risky activity beats the safe one in the current flow if and only if $p\rho > q$. We call $p = q/\rho$ the *myopic curve*. It is the locus at which a player who ignores the future is indifferent. Above it, the expected flow from experimentation exceeds the safe flow. Our figures and numerical examples use the *baseline parameters* $N = 3$, $\lambda = 1$, $h = 2$, $s = 1$, $\eta = 1$ and $r = 0.2$, for which $q = 1/2$.

Two externalities. An opportunity found by one player shows everybody that the activity is productive, which raises every player's continuation value, and it resets quality for everybody. Players can free ride on each other's experiments, and each wants to exploit the resource before the others deplete it. On the edge $p = 1$ nothing is left to learn and only depletion matters. Before a discovery, information matters most at full quality, where the decision to stop experimenting is effectively taken. Figure 2 shows how play unfolds in the equilibrium of Section 4, when the resource is productive and when it is not.

2.2 Strategies and equilibrium

A *Markov strategy* is a Borel function $a : [0, 1]^2 \to [0, 1]$. A player i who uses the strategy a_i exerts effort $a_t^i = a_i(p_t, \rho_t)$. A strategy profile is *symmetric* if all players use the same strategy.

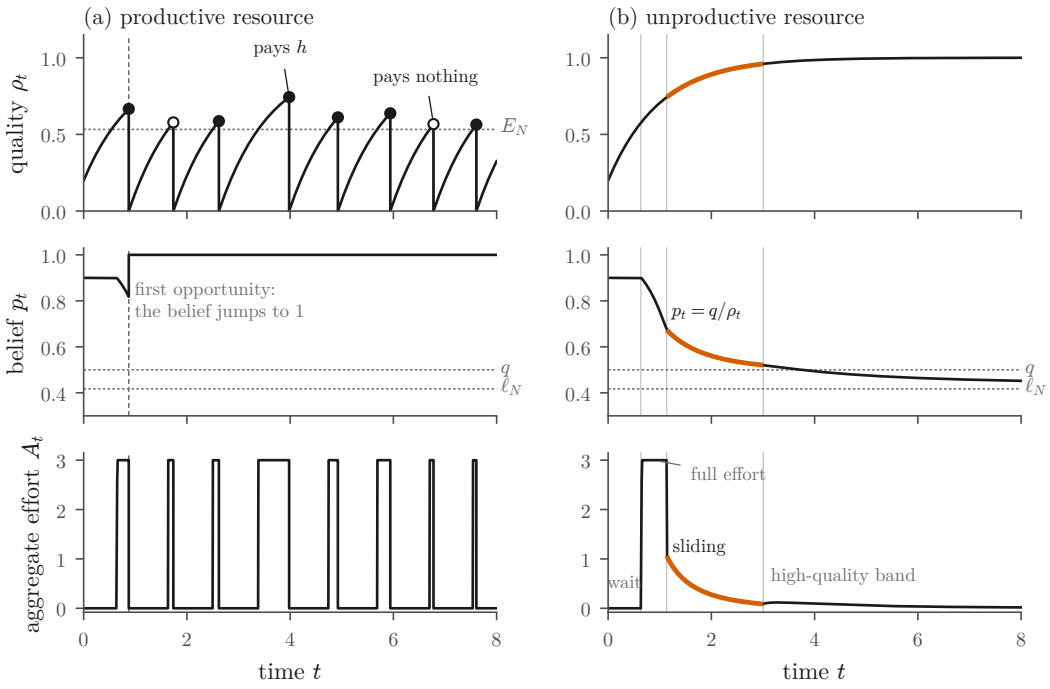


Figure 2: Quality recovers between opportunities and drops to zero at each one, while the belief falls until the first opportunity shows that the resource is productive. Equilibrium play from belief $p_0 = 0.9$ and quality $\rho_0 = 0.2$, when the resource is productive (a) and when it is not (b). Top: quality ρ_t, with opportunities that pay the prize h (filled circles) or nothing (open circles). Middle: the players' common belief p_t that the resource is productive. Bottom: aggregate effort A_t of the three players. (a) The first opportunity arrives at its median date given that the resource is productive. Afterwards players wait until quality reaches $E_N = 0.53$ and then use full effort until the next opportunity, whose dates and outcomes are drawn with a fixed seed. (b) No opportunity arrives. Players wait, use full effort, slide along the myopic curve $p = q/\rho$, on which working the resource pays exactly the safe flow (vermilion), and then experiment with effort that vanishes as the belief approaches $\ell_N = 0.42$. Baseline parameters, as in Figure 1.

Markov strategies can jump. The equilibrium we construct switches from waiting to full effort on the myopic curve at intermediate qualities. Where the opponents' strategies are continuous, the belief moves according to (2.2) with the efforts that the strategies prescribe. On a curve where they jump, we follow Filippov (1988) and let the opponents' aggregate effort take any value between its limits on the two sides of the curve. If effort on both sides pushes the state toward the curve, the state then slides along it, with the aggregate effort that keeps it there. Ordinary trajectories are not enough here, because for regular symmetric profiles at intermediate qualities some deviations have none (Proposition 5). For the equilibrium we construct, the convention describes the limit of play when players revise their effort at frequent discrete dates (Proposition 6).

A *unilateral deviation* is any effort process adapted to the public history and to independent private randomization. Before the first opportunity, and conditional on the deviator's private randomization, a deviation is a measurable effort schedule u along the no-discovery trajectory. After the first opportunity, it may be arbitrary.

When all players follow a symmetric strategy, we require the state to move according to the actions that the strategy prescribes, including its actions on any curve where it jumps, and we call the resulting path its *profile path*. A profile is *well posed against unilateral deviations* if every unilateral deviation, from every initial state, generates a unique trajectory under this convention. Section 7.1 defines both notions formally.

Definition 1. A symmetric Markov strategy a^* is a *symmetric Markov perfect equilibrium under the Filippov convention* if, from every initial state, its profile path exists and is unique; the profile is well posed against unilateral deviations; and, for every initial state $(p_0, \rho_0) \in [0,1]^2$ and every player i, no unilateral deviation yields player i a strictly higher payoff (2.3) than a^*.

3 Two benchmarks

3.1 Competition after discovery

After a discovery the state is known forever and only quality moves. Recovery is independent of effort until the next opportunity, so the edge $p = 1$ is a one-dimensional game in ρ. Put

$$\Lambda_N = r + \lambda N, \qquad L_N = \Lambda_N + \eta.$$

Once the resource is known to be productive, players wait until quality reaches a threshold E_N and then use full effort until the next opportunity.

Proposition 1 (Known-state equilibrium)**.** *There is a unique $E_N \in (q, 1)$ solving*

$$q = E_N - \frac{\lambda\eta}{L_N} \frac{(1-E_N)\big[1-(1-E_N)^{\beta}\big]}{r + \lambda K\big[1-(1-E_N)^{\beta}\big]}. \tag{3.1}$$

With $z_N = 1 - E_N$ and $H_N = 1 - z_N^{\beta}$, put

$$e_N = \frac{r\eta g}{L_N} \frac{z_N^{1+\beta}}{r + \lambda K H_N} > 0. \tag{3.2}$$

The profile in which all players wait when $\rho < E_N$ and use full effort when $\rho > E_N$ is a Markov perfect equilibrium of the game on $\{p = 1\}$, optimal against every unilateral deviation, with value

$$u(\rho) = \begin{cases} s + e_N(1-\rho)^{-\beta}, & \rho \le E_N, \\ \dfrac{\lambda N}{\Lambda_N}(s + e_N) + \dfrac{\eta r g}{\Lambda_N L_N} + \dfrac{rg}{L_N}\rho, & \rho \ge E_N. \end{cases} \tag{3.3}$$

In particular, the value immediately after an opportunity is $u(0) = s + e_N$.

The constant e_N is the surplus each player obtains from the known resource. A discovery affects future payoffs only through e_N, which summarizes the post-discovery game in everything that

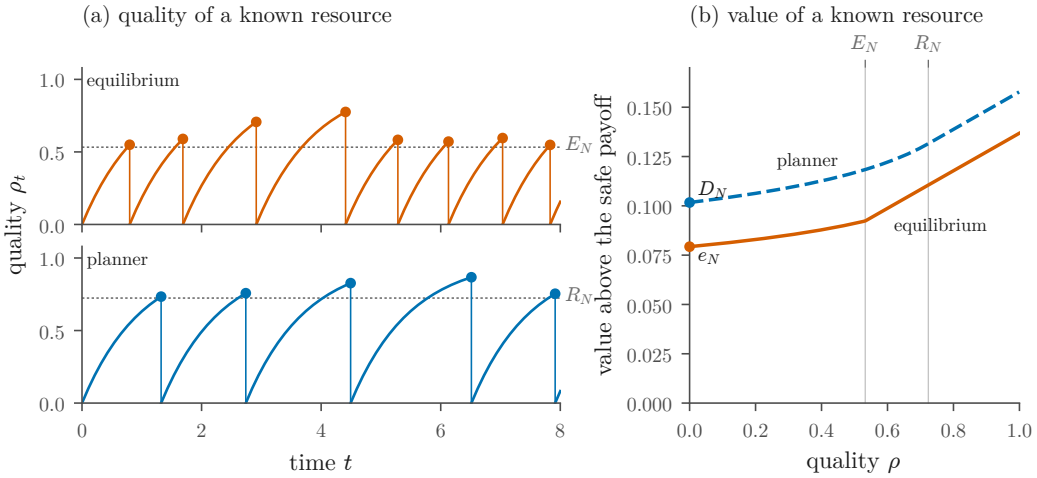


Figure 3: Once the resource is known to be productive, players harvest it too early. (a) Quality over time when players start exploiting as soon as quality reaches $E_N = 0.53$, as in equilibrium, and when they wait until $R_N = 0.72$, as the planner would. Both use full effort once they exploit, and both rows use the same random lengths of the full-effort spells. Dots mark opportunities, each of which pays the prize with probability equal to the quality at that date. In expectation an opportunity arrives every 1.09 units of time at quality 0.65 in equilibrium, and every 1.62 units at quality 0.79 under the planner. (b) The value of a known productive resource per player, above the value of the safe activity, as a function of quality. Right after an opportunity, when quality is zero, it is $e_N = 0.079$ in equilibrium and $D_N = 0.102$ under the planner. Baseline parameters, as in Figure 1.

follows. Players start exploiting the known resource at quality E_N, below the planner's known-state threshold R_N characterized next, so part of the resource is wasted.

3.2 Efficient experimentation

The planner chooses aggregate effort $A \in [0, N]$ to maximize the players' average payoff, and we write $V(p, \rho)$ for the planner's value per player. In the known state the planner waits below a threshold $R_N \in (q, 1)$ and uses full effort above it. The threshold and the planner's surplus per player after a discovery, $D_N = V(1, 0) - s$, are

$$q = R_N - \frac{\lambda N \eta}{r L_N}(1 - R_N)\left[1 - (1 - R_N)^{\beta}\right], \qquad D_N = \frac{\eta g}{L_N}(1 - R_N)^{1+\beta}. \tag{3.4}$$

Because the planner internalizes depletion, $E_N < R_N$ and $e_N < D_N$ (Section A). Figure 3 shows that equilibrium players harvest a known resource more often and at lower quality than the planner would.

For uncertain beliefs, define

$$\Pi_N = \lambda(N D_N + rh) - rs > 0, \qquad \Omega = g(r + \eta), \qquad p_N^P = \frac{rs}{rs + \Pi_N} = \frac{rs}{\lambda(N D_N + rh)}.$$

We call p_N^P the *planner's abandonment belief*. The *planner's frontier* is a decreasing curve $p = p^P(\rho)$

on $[R_N, 1]$ with two branches that meet at a junction (ρ_c^P, p_c^P) given in closed form in Section A. On the high-quality branch, $\rho_c^P \le \rho \le 1$,

$$p^P(\rho) = \frac{rs}{rs + \Pi_N - \Omega(1-\rho)}, \qquad A^s(\rho) = \frac{\eta\,\Omega(1-\rho)}{\lambda[\Pi_N - \Omega(1-\rho)]}, \tag{3.5}$$

and on the low-quality branch the frontier is given parametrically by (A.5).

Theorem 1 (Efficient experimentation)**.** *The planner's problem has a solution from every initial state, and the following stationary policy is optimal. At $p = 0$ use zero effort. On $\{p = 1\}$, use full effort if $\rho > R_N$ and wait otherwise. For $0 < p < 1$, wait if $\rho \le R_N$ or $p < p^P(\rho)$; use full effort if $p > p^P(\rho)$ and also on the low-quality branch $p = p^P(\rho)$, $R_N < \rho < \rho_c^P$; on the high-quality branch $p = p^P(\rho)$, $\rho \ge \rho_c^P$, use aggregate effort $A^s(\rho) \in [0, N]$, which keeps the state on the frontier. For an uncertain prior the optimal no-discovery path is unique; it consists of waiting, full effort and declining effort on the frontier, in this order, and any phase may be empty. Its limiting belief is $\min\{p_0, p_N^P\}$, and for $0 < p_0 < 1$:*

$$\int_0^\infty A_t^P \,\mathrm{d}t = \frac{1}{\lambda}\big[\mathrm{logit}(p_0) - \mathrm{logit}(p_N^P)\big]_+, \qquad \mathbb{P}(\tau < \infty) = \frac{(p_0 - p_N^P)_+}{1 - p_N^P}, \tag{3.6}$$

where A_t^P is aggregate effort along the optimal no-discovery path and τ is the date of the first opportunity. These quantities do not depend on the initial quality ρ_0.

Figure 4 shows the policy and three optimal paths. On the high-quality branch, declining effort keeps beliefs and quality on the frontier, and the posterior converges to p_N^P as quality approaches one. This is a singular arc of the planner's control problem.

The proof writes the planner's objective in terms of the probability $Q_t = \exp(-\lambda \int_0^t A)$ of no opportunity by t conditional on productivity. For an uncertain prior the objective is strictly concave in Q (Lemma A.1), and a multiplier argument identifies and verifies its unique maximizer.

4 Markov perfect equilibrium

4.1 The shape of the equilibrium

After a discovery all players follow the known-state equilibrium of Proposition 1. Before a discovery, the equilibrium we construct is organized by two thresholds of quality above E_N,

$$\rho_B = 1 - \frac{e_N}{\eta h}, \qquad \rho_* = \frac{\eta\rho_B + \lambda K q}{\eta + \lambda K}. \tag{4.1}$$

These thresholds satisfy $q < E_N < \rho_* < \rho_B < 1$ for all primitives (Lemma B.2). Figure 5 shows the equilibrium at the baseline parameters, and Table 1 summarizes it.

Below E_N nobody experiments, whatever the belief, because even a resource known to be productive is not yet worth exploiting. Between E_N and ρ_*, players wait below a lower frontier

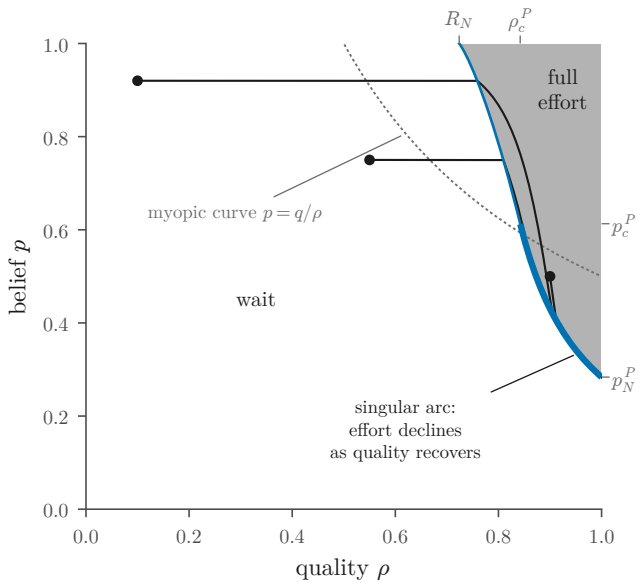


Figure 4: The planner waits for quality to recover, experiments with full effort, and finally rides a curve on which effort declines as quality recovers. The planner waits in the white region and uses full effort in the grey region. Its frontier (blue) has two branches that meet at (ρ_c^P, p_c^P). On the thin branch the planner switches to full effort, and the thick branch is a singular arc, on which aggregate effort $A^s(\rho)$ keeps the state on the curve. The frontier meets the edge $p = 1$ at R_N, the quality at which the planner starts exploiting a resource known to be productive, and the edge $\rho = 1$ at p_N^P, the planner's abandonment belief, at which efficient learning stops. Black curves are optimal paths without a discovery from three initial states (dots). The dotted curve is the myopic curve $p = q/\rho$, on which working the resource pays exactly the safe flow. Baseline parameters, as in Figure 1.

p_L, use full effort above an upper frontier p_H, and use interior effort in a narrow band between the two, above the myopic curve. Both frontiers start at $p = 1$ when quality is E_N, so just above E_N players experiment only if they are nearly certain that the resource is productive. Between ρ_* and ρ_B the band closes, and players wait below the myopic curve $p = q/\rho$ and use full effort above it. Above ρ_B a band of interior effort opens below the myopic curve. At full quality this band is the symmetric equilibrium of Keller et al. (2005), with the value e_N of the known resource as the continuation after a success, and effort in it rises from zero at the *individual abandonment belief* $\ell_N = rs/[\lambda(rh + e_N)]$ to one at the *upper full-quality cutoff* $\bar{p}_N$. A single player who experimented alone at full quality, with continuation e_N after a success, would stop at ℓ_N.

Between ρ_* and ρ_B the switch on the myopic curve attracts the state from both sides (Figure 6a). Below the curve, players wait, and recovery lowers the myopic curve toward the state. Above it, full effort lowers the posterior faster than the curve falls. Once on the curve, the state slides along it with aggregate effort

$$A_m(\rho) = \frac{\eta(1-\rho)}{\lambda(\rho - q)}, \tag{4.2}$$

which makes the posterior fall exactly as fast as the myopic curve. We call this part of the myopic curve the *sliding arc*. Aggregate effort on the arc does not depend on the number of players, and on

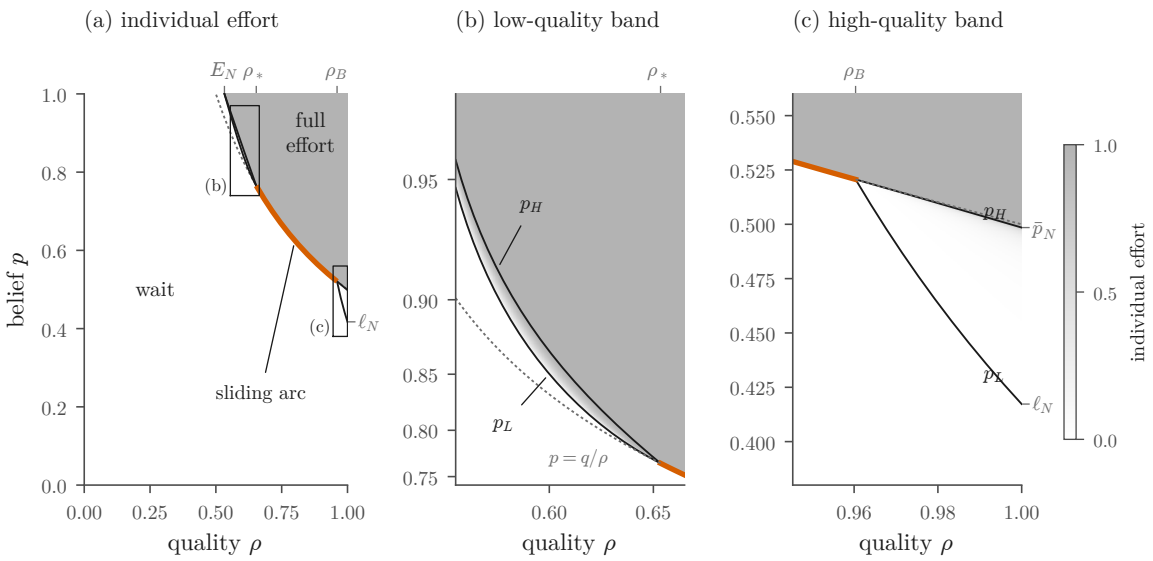


Figure 5: In equilibrium, players wait at low quality, use full effort at high beliefs, and slide along the myopic curve at intermediate qualities. Shading shows each player's effort, from white (wait) to grey (full effort). (a) The whole state space. Black curves are the frontiers of the two bands of interior effort, the vermilion curve is the sliding arc on the myopic curve $p = q/\rho$ (dotted elsewhere), and the rectangles mark the regions enlarged in (b) and (c). On the top edge, E_N is the quality at which players start exploiting a resource known to be productive, and the sliding arc runs from ρ_* to ρ_B. On the right edge, ℓ_N is the belief at which equilibrium learning stops. (b) The low-quality band between the frontiers p_L and p_H, at qualities between E_N and ρ_*, on a logit scale. (c) The high-quality band between p_L and p_H above ρ_B. At full quality it runs from ℓ_N to $\bar{p}_N$. Baseline parameters, as in Figure 1.

the arc each player's flow payoff equals the safe flow s (Theorem 2).

A player on the myopic curve who postpones a unit of effort for a moment loses nothing in the current flow, but she trades off two effects. Quality recovers in the meantime, which raises the expected prize of an opportunity at rate $\eta h(1-\rho)$, and a discovery, worth e_N to the player, comes later. The first effect dominates exactly when $\rho < \rho_B$. Below ρ_B, nobody experiments at a flow loss for the sake of information. Just above the curve, the other players use full effort, so if the resource is productive one of them is likely to find the next opportunity first, take its prize and reset quality. This threat outweighs the gain from waiting for recovery exactly when $\rho > \rho_*$, so above ρ_* the switch to full effort happens on the curve itself.

On the myopic curve $gp\rho = s$, so a player earns exactly the safe flow whatever effort the player exerts. Each player wants to take the surplus above the curve before the others deplete it, and nobody wants to experiment at a flow loss below the curve when the others will experiment there once quality has recovered. Aggregate effort keeps the state on the curve and dissipates the flow rent, much as entry drives the rent of an open-access fishery to zero (Gordon, 1954). Here the posterior plays the role of the fishery's stock, and aggregate effort $A_m(\rho)$ falls as quality recovers (Figure 6c).

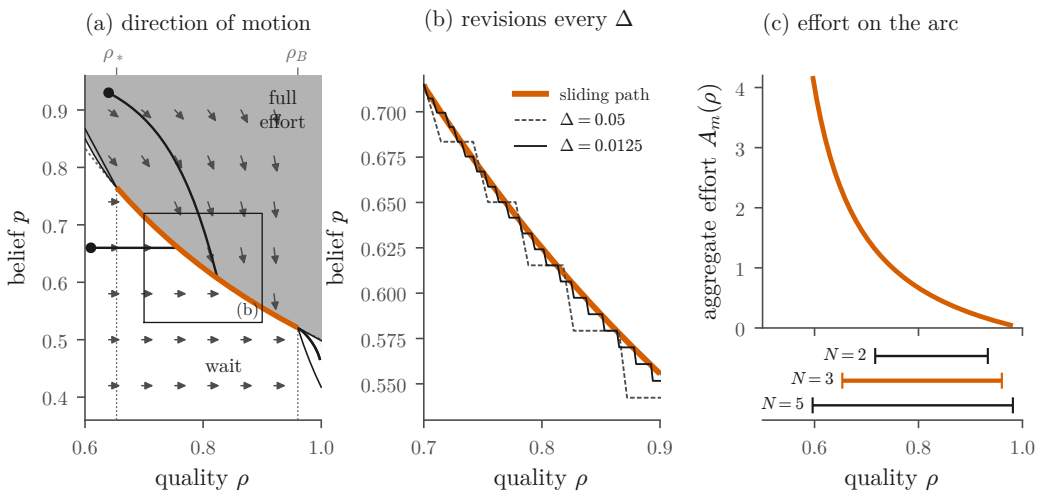


Figure 6: At intermediate qualities, effort on both sides pushes the state onto the curve on which working the resource pays exactly the safe return, and the state slides along it. (a) Arrows give the direction of motion under equilibrium effort, not its speed. Below the dotted myopic curve $p = q/\rho$ players wait (white) and only quality moves. Above it they use full effort (grey), and the belief falls faster than the curve. Between $\rho_* = 0.65$ and $\rho_B = 0.96$ both directions point at the curve, and two paths (black) reach it and slide along the sliding arc (vermilion). (b) The rectangle marked in (a). If players revise their effort only every Δ units of time, the state zigzags around the curve, and the zigzag shrinks with Δ toward the sliding path. (c) Top: aggregate effort on the arc, $A_m(\rho) = \eta(1-\rho)/[\lambda(\rho - q)]$, is the effort that keeps the state on the curve while quality recovers, whatever the number of players N. Bottom: the range of qualities on which players slide, for $N = 2$, 3 and 5. Baseline parameters otherwise, as in Figure 1.

4.2 Existence

The construction requires three conditions on primitives, which Section 6.3 states formally as Condition 1. Condition (R1) connects the upper frontier of the low-quality band to the onset of sliding, (R2) makes full-effort paths enter the high-quality band, and (R3) keeps the low-quality band above the myopic curve. Proposition 2 gives sufficient conditions on primitives, and the three conditions hold for all sufficiently large N.

Proposition 2 (Sufficient conditions). *(a)* (R1) *holds if* $r \le K(\lambda K + \eta)$, *in particular if* $r \le \lambda(N-1)$.

(b) (R2) *holds if* $\eta^2 \le \lambda K\big[Nr + K(\lambda N + \eta)\big]$, *in particular if* $\eta \le \lambda(N-1)$.

(c) (R3) *holds if both of the following hold:*

$$\text{(R3a)} \quad \log\frac{\rho_B - q}{E_N - q} < \frac{\lambda N}{\eta}\log\frac{1-E_N}{1-\rho_B};$$

$$\text{(R3b)} \quad (1-E_N)^{\beta} \le \frac{\lambda}{\lambda + r} \quad \textit{or} \quad \frac{1-\rho_\varepsilon}{1-E_N} \ge M(t_* - t_E),$$

where $\rho_\varepsilon = q + e_N/g$, $M(\tau) = \Lambda_N(1 - e^{-L_N\tau})/[L_N(1 - e^{-\Lambda_N\tau})]$, *and* $t_* - t_E$ *is the time that*

Table 1: The symmetric Markov perfect equilibrium. On the sliding arc, aggregate effort $A_m(\rho) = \eta(1-\rho)/[\lambda(\rho-q)]$ keeps the state on the myopic curve. The frontiers and the interior efforts are defined in Section 6.3.

Quality	Beliefs	Individual effort
$\rho \le E_N$	$0 \le p < 1$	0 (wait)
$E_N < \rho < \rho_*$	$p \le p_L(\rho)$	0
	$p_L(\rho) < p < p_H(\rho)$	$\mathcal{N}/(\lambda K m) \in (0,1)$, increasing in p
	$p \ge p_H(\rho)$	1
$\rho_* \le \rho \le \rho_B$	$p < q/\rho$	0
	$p = q/\rho$	$A_m(\rho)/N$; the state slides along the curve
	$p > q/\rho$	1
$\rho_B < \rho \le 1$	$p \le p_L(\rho)$	0
	$p_L(\rho) < p < p_H(\rho)$	$\alpha(p,\rho)$ in (6.10), increasing in p
	$p \ge p_H(\rho)$	1
any	$p = 1$	known-state equilibrium (Proposition 1)

quality takes to recover from E_N to ρ_*.

(d) For all other primitives fixed, (R1)–(R3) *hold for all sufficiently large N.*

At the baseline parameters the three conditions hold (Lemma D.3). On a 33×33 grid of primitives with $q = 1/2$, numerical tests of all three pass in 76%, 89% and 97% of cells for $N = 2$, 3 and 5 (Figure 14).

The frontiers $p_L \le p_H$ and the interior efforts in Table 1 are defined in Section 6.3, and the private gain in the theorem below is a player's marginal value of own effort, defined in (6.2).

Theorem 2 (Markov perfect equilibrium)**.** *Assume Condition 1. The symmetric Markov strategy a^* described in Table 1, with individual effort A_m/N on the sliding arc, is a symmetric Markov perfect equilibrium under the Filippov convention (Definition 1), and it has the following properties.*

(a) The frontiers $p_L \le p_H$ are continuous and strictly decreasing on $(E_N, 1]$; both tend to one as $\rho \downarrow E_N$; they coincide exactly on $[\rho_*, \rho_B]$, where they equal the myopic curve; and $p_L(1) = \ell_N < \bar{p}_N = p_H(1) < q$.

(b) The strategy a^* is nondecreasing in p. For each ρ it is continuous in $p \in (0,1)$, except at $p = q/\rho$ when $\rho_* \le \rho \le \rho_B$, where it jumps from zero to one; as a function of (p,ρ) it is continuous on $(0,1) \times [0,1]$ off the closed sliding arc.

(c) The equilibrium value v^* is continuous on $[0,1]^2$ and continuously differentiable on $(0,1)\times(0,1)$. It equals s for $p \le \ell_N$, and the known-state value $u(\rho)$ at $p = 1$. At uncertain interior states the private gain is negative in the interior of the waiting region, zero on the closed bands and the arc, and positive in the interior of the full-effort region.

(d) On the sliding arc each player's flow payoff equals the safe flow s, and aggregate effort $A_m(\rho)$ does not depend on N.

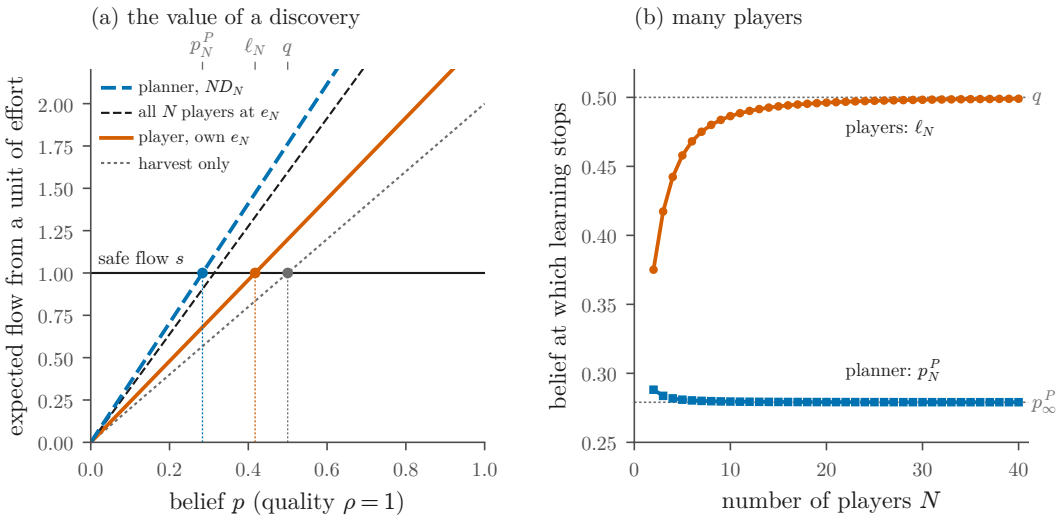


Figure 7: Players stop learning at a higher belief than the planner because a discovery is worth less to them. (a) At full quality, a unit of effort yields the expected harvest gp, with $g = \lambda h$, plus the value of a possible discovery, which arrives at rate λp. In flow terms the discovery is worth $(\lambda/r)pe_N$ to a player, who counts only her own surplus e_N from a known resource, and $(\lambda/r)pND_N$ to the planner, who counts the surplus $D_N > e_N$ of all N players. Experimenting beats the safe flow s above $\ell_N = 0.42$ for a player and above $p_N^P = 0.28$ for the planner, and a player who ignored discoveries would stop at $q = 0.5$. The thin dashed line counts all N players at the equilibrium surplus e_N, which shows that most of the gap comes from counting all players. (b) As the number of players grows, competition dissipates each player's surplus from a discovery. Equilibrium learning then stops at q, while the planner's abandonment belief converges to $p_\infty^P < q$. Baseline parameters otherwise, as in Figure 1.

5 Too early and too little

In this section t denotes calendar time, measured from the initial date, and τ the date of the first opportunity.

5.1 The two distortions

The planner's abandonment belief $p_N^P = rs/[\lambda(ND_N + rh)]$ differs from $\ell_N = rs/[\lambda(rh + e_N)]$ because the planner values a discovery for all N players and values the known resource at $D_N > e_N$. Both differences lower the abandonment belief, so $p_N^P < \ell_N$ (Figure 7a). On the known-state edge the comparison goes the other way, and the equilibrium starts exploiting the resource at $E_N < R_N$.

Theorem 3 (Too early and too little)**.** *Assume Condition 1, and let A^* and A^P denote equilibrium and efficient aggregate effort.*

(a) For every $\rho \in (E_N, R_N)$ and every $p \in (p_H(\rho), 1]$, $A^*(p, \rho) = N$ and $A^P(p, \rho) = 0$.

(b) From every initial state with $0 < p_0 < 1$, along the no-discovery path of the equilibrium of Theorem 2 the posterior converges to $p_\infty = \min\{p_0, \ell_N\}$, and, with A_t^* the aggregate effort at

date t along the path, total effort and the probability of a discovery are

$$\int_0^\infty A_t^* \,\mathrm{d}t = \frac{1}{\lambda}\big[\operatorname{logit}(p_0) - \operatorname{logit}(\ell_N)\big]_+, \qquad \mathbb{P}(\tau < \infty) = \frac{(p_0 - \ell_N)_+}{1 - \ell_N}. \tag{5.1}$$

None of these quantities depends on the initial quality ρ_0. If $p_0 > \ell_N$, then $p_t > \ell_N$ for all t and $\int_T^\infty A_t^* \,\mathrm{d}t > 0$ for every T, so experimentation never stops and a discovery occurs after any given date with positive probability.

(c) *For every $p \in (p_N^P, \ell_N)$ there is $\bar\rho(p) < 1$ such that, for every $\rho > \bar\rho(p)$, $A^P(p,\rho) = N$ and the equilibrium never experiments from (p,ρ). From every initial state with $p_N^P < p_0 < 1$, total efficient effort exceeds total equilibrium effort by $\big[\operatorname{logit}\min\{p_0, \ell_N\} - \operatorname{logit} p_N^P\big]/\lambda > 0$.*

(d) *From every initial state with $0 < p_0 < 1$, the no-discovery path of the equilibrium visits waiting, the low-quality band, full effort, the sliding arc and the high-quality band in this order; any of these phases may be empty, and the high-quality band, once entered, is never left.*

Figure 1 shows the distortions. When productivity is likely, information is worth little and depletion is the main concern, so players exploit a resource that the planner would let recover. Free riding matters most at high quality and pessimistic beliefs. At full quality, equilibrium effort falls short of efficient effort on $(p_N^P, \bar p_N)$ and is zero on $(p_N^P, \ell_N]$. Around the sliding arc, the effort comparison can go either way, depending on how the myopic curve and the planner's frontier cross.

For $p_0 > \ell_N$ the posterior approaches ℓ_N because, as quality approaches one, the equilibrium approaches the band at full quality. The lower frontier ℓ_N of that band is the belief at which a single player is indifferent between the safe flow s and the expected flow from experimenting, $gp + (\lambda/r)pe_N$, which includes the value e_N of a discovery. In the band, effort vanishes quadratically in the distance from ℓ_N, and the resulting long tail of small experiments is described by Remark 1.

Remark 1 (The tail of discovery). Assume Condition 1, fix $\ell_N < p_0 < 1$, and put

$$y_{m,\infty} = \operatorname{logit} q - \operatorname{logit} \ell_N, \qquad c_N = \frac{Nr}{2K(1 - e^{-y_{m,\infty}})}.$$

From every initial quality, along the no-discovery path,

$$\operatorname{logit} p_t - \operatorname{logit} \ell_N \sim \frac{1}{c_N t}, \qquad \mathbb{P}(t < \tau < \infty) \sim \frac{(1-p_0)\ell_N}{(1-\ell_N)c_N t}. \tag{5.2}$$

Consequently $\mathbb{E}[\tau \mid \tau < \infty] = \infty$ for every finite N.

The conditional mean is infinite because of rare discoveries after long spells of very low effort, even though total experimentation is finite.

Figure 8 compares equilibrium and efficient paths from two initial states. From a confident start, players use full effort at a lower quality than the planner, and from either start the planner ends at a lower belief. Figure 9 illustrates parts (b) and (c) of Theorem 3. Paths from the same prior and different initial qualities end at the same belief, and from a prior of 0.5 a productive resource is

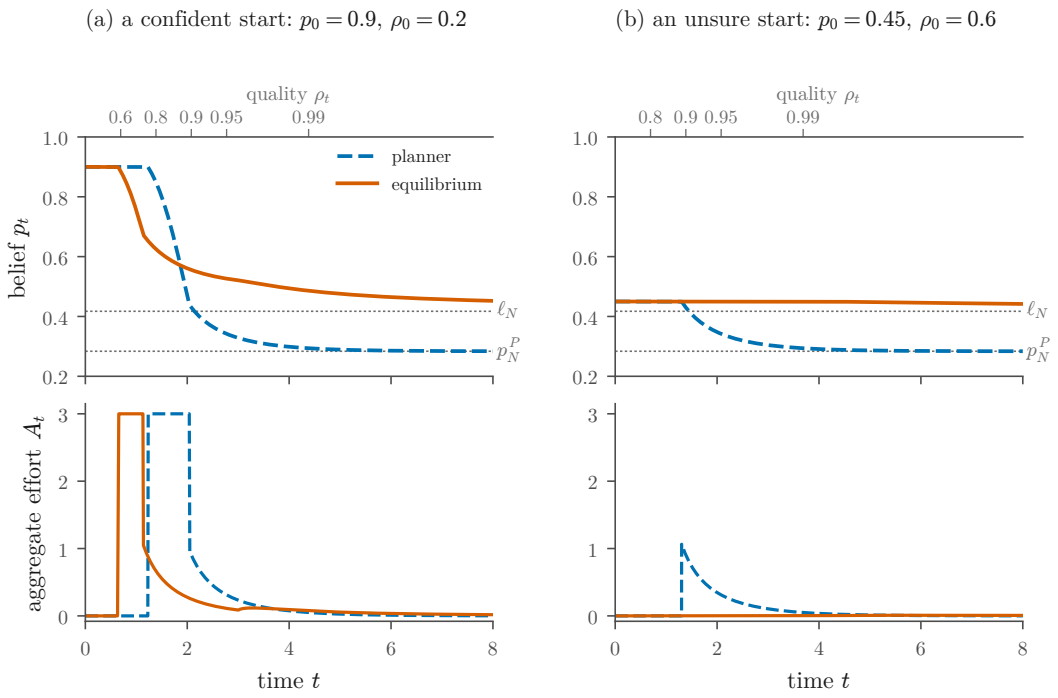


Figure 8: Players start experimenting at a lower quality than the planner and stop learning at a higher belief. Belief p_t (top) and aggregate effort A_t (bottom) along the paths without a discovery of the equilibrium (vermilion) and of the planner (dashed blue) from the same initial state. Quality ρ_t, shown on the top axis, rises in the same way along both paths. (a) From a confident start, $p_0 = 0.9$ and $\rho_0 = 0.2$, players use full effort earlier and then slide, while the planner waits longer, experiments longer and ends at a lower belief. (b) From an unsure start, $p_0 = 0.45$ and $\rho_0 = 0.6$, players barely experiment, while the planner experiments as quality approaches one. ℓ_N and p_N^P are the beliefs at which equilibrium and efficient learning stop. Baseline parameters, as in Figure 1.

found with probability 0.28 in equilibrium and 0.60 under the planner.

The belief at which learning stops is not specific to the equilibrium we construct. In every Nash equilibrium in which players commit to deterministic effort schedules until the first opportunity and follow the known-state equilibrium afterward, including asymmetric ones, learning stops at $\min\{p_0, \ell_N\}$ (Theorem 6). At full quality and from priors above ℓ_N, such an equilibrium reaches ℓ_N in finite time and gives every player a strictly higher payoff than the Markov perfect equilibrium, in which indifference in the band slows experimentation down (Section F).

5.2 Welfare

The planner's value exceeds the equilibrium value by at least the discounted gain from using the known resource efficiently after a discovery, and the gap is positive exactly at beliefs above the planner's abandonment belief.

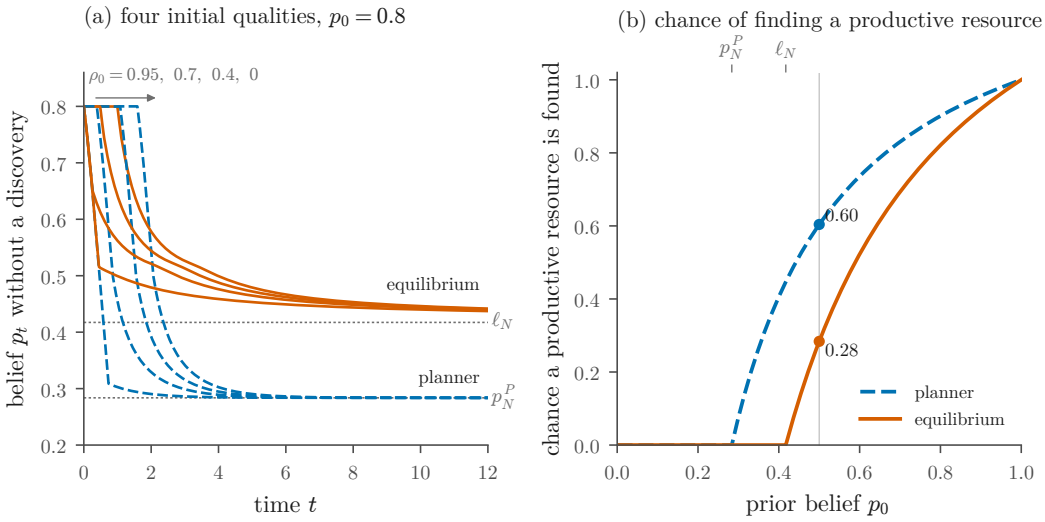


Figure 9: How much is learned does not depend on initial quality, and the planner learns more. (a) The belief without a discovery from the prior $p_0 = 0.8$ and four initial qualities, $\rho_0 = 0.95$, 0.7, 0.4 and 0. Paths that start from lower quality move later. All equilibrium paths (vermilion) converge to ℓ_N, and all the planner's paths (dashed blue) to p_N^P. (b) The probability that a productive resource is ever found, as a function of the prior, which does not depend on initial quality. From a prior of 0.5, it is 0.28 in equilibrium and 0.60 under the planner. Baseline parameters, as in Figure 1.

Proposition 3 (Welfare)**.** *Assume Condition 1. For every state,*

$$V(p,\rho) - v^*(p,\rho) \ \geq\ \mathbb{E}\big[e^{-r\tau}\mathbf{1}_{\{\tau<\infty\}}\big](D_N - e_N) \ \geq 0, \tag{5.3}$$

where τ is the date of the first opportunity on the equilibrium path. The planner does strictly better, $V(p,\rho) > v^*(p,\rho)$, if and only if $p > p_N^P$.

The bound comes from a policy available to the planner, who can replicate equilibrium aggregate effort until the first opportunity, including the effort that keeps the state on the sliding arc, and then use the known resource efficiently. The bound leaves out the planner's further gain from better effort before a discovery, which in this model means experimenting later when productivity is likely and longer when quality is high. Figure 10 shows the values at full quality and the welfare loss at three qualities.

5.3 Large groups

As the number of players grows, competition for the known resource intensifies and the surplus it yields to each player vanishes quickly, with $e_N = C_E/N^2 + o(N^{-2})$ and $C_E = r\eta h(1-q)^{1+\beta}/\{\lambda[1-(1-q)^{\beta}]\}$. The value of a discovery to an individual player therefore vanishes, and so does each player's private incentive to experiment for information. The planner values a discovery at ND_N, which converges to a positive limit.

Proposition 4 (Many players)**.** *Fix r, λ, h, s, η and let $N \to \infty$.*

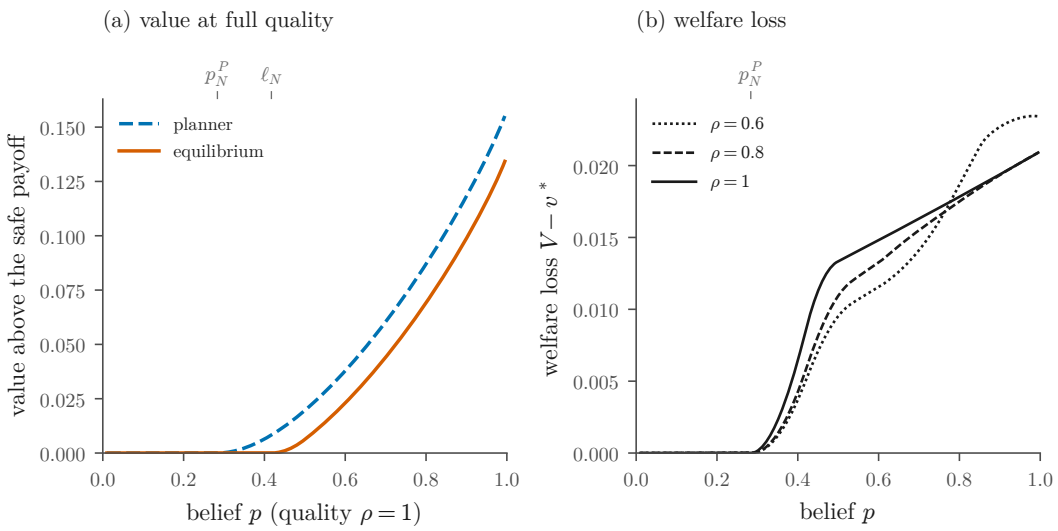


Figure 10: The planner does better at every belief above its abandonment belief p_N^P. (a) Values at full quality, above the value of the safe activity, as functions of the belief. The equilibrium value is above the safe value only above ℓ_N, and the planner's above p_N^P. (b) The welfare loss, the planner's value V minus the equilibrium value v^*, at three qualities. It is zero at beliefs below p_N^P and positive above. Baseline parameters, as in Figure 1.

(a) The quality thresholds satisfy $E_N - q \sim \eta(1-q)/(\lambda N^2)$, $\rho_* - q \sim \eta(1-q)/(\lambda N)$ *and* $1 - \rho_B \sim C_E/(\eta h N^2)$. *The individual abandonment belief and the upper full-quality cutoff satisfy*

$$q - \ell_N \sim \frac{qC_E}{rh}\, N^{-2}, \qquad q - \bar{p}_N \sim \frac{qC_E^2}{2\lambda r h^2 (1-q)}\, N^{-5}.$$

(b) Condition 1 holds for all sufficiently large N.

(c) For every compact set $\mathcal{C} \subset \{(p,\rho) : 0 < p < 1,\ q < \rho < 1,\ p \neq q/\rho\}$ *there is* $N_{\mathcal{C}}$ *such that, for* $N \geq N_{\mathcal{C}}$, *the equilibrium of Theorem 2 uses zero effort at the points of* $\mathcal{C}$ *below the myopic curve and full effort at those above it. On the myopic curve, at qualities in* $[\rho_*, \rho_B]$, *an interval that expands to* $(q, 1)$, *the state slides with aggregate effort* $A_m(\rho)$, *which does not depend on* N.

(d) From any $0 < p_0 < 1$, *total equilibrium effort converges to* $[\mathrm{logit}(p_0) - \mathrm{logit}(q)]_+/\lambda$, *whereas total efficient effort converges to* $[\mathrm{logit}(p_0) - \mathrm{logit}(p_\infty^P)]_+/\lambda$ *with* $p_\infty^P = rs/[\lambda(C_P + rh)] < q$, *where* $C_P = \lim_N ND_N > 0$.

On the sliding arc, recovery pins down aggregate effort $A_m(\rho)$, so adding players lowers each player's share $A_m(\rho)/N$ and expands the range of qualities on which sliding occurs. The planner's abandonment belief stays bounded away from q, so the shortfall in learning persists in large groups.

For the equilibrium constructed here, the terminal belief, the probability of a discovery and total effort are within $O(N^{-2})$ of their limits, uniformly over initial beliefs and qualities (Lemma D.2). At full quality the upper full-quality cutoff is within order N^{-5} of q, the belief at which the myopic

curve meets full quality, so full effort ends almost exactly at q and is followed by a thin band of declining effort. Figure 6(c) shows the range of qualities on which players slide for $N = 2$, 3 and 5, and Figure 7(b) shows where learning stops as N grows.

6 Construction of the equilibrium

A change of units makes the game before a discovery deterministic, and in the resulting game each player's private gain from effort satisfies a transport equation along paths of constant effort.

6.1 A deterministic reduction

We keep the known-state continuation of Proposition 1 fixed throughout. Before a discovery, it is convenient to measure quality by the *quality clock*

$$t = -\frac{1}{\eta}\log(1-\rho), \qquad \rho(t) = 1 - e^{-\eta t},$$

which advances one-for-one with calendar time between opportunities, and the belief by its log odds $x = \operatorname{logit} p$. From now on we also index time by the quality clock, so that the no-discovery state moves according to $\dot{x} = -\lambda A$, $\dot{t} = 1$. Define the *myopic surplus*

$$m(x,t) = e^{x}\big(g\rho(t) - s\big) - s = \frac{gp\rho - s}{1-p},$$

which, for $\rho(t) > q$, is positive above the myopic curve $x_m(t) = \log q - \log(\rho(t) - q)$ and negative below it. For $\rho(t) \le q$ it is negative everywhere.

Lemma 6.1 (Reduction)**.** *Fix a well-posed profile, an initial state (p,ρ) with $0 < p < 1$ and $\rho < 1$, and a deviation by player i with no-discovery schedule u and trajectory (x_t), starting at (x,t), along which the opponents' aggregate effective effort is A_{-i} (Section 7.1). If player i follows the known-state equilibrium after a discovery, then $J^i - s = (1-p)\,W^i$, where*

$$W^i = \int_t^\infty e^{-r(t'-t)}\Big[\, r\,u_{t'}\,m(x_{t'},t') + \lambda e_N\, e^{x_{t'}}\big(u_{t'} + A_{-i,t'}\big)\Big]\,\mathrm{d}t'. \tag{6.1}$$

If player i deviates after a discovery as well, then $J^i - s \le (1-p)W^i$. The reduction rests on an identity of first discoveries. With $Q_{t'} = \exp\{-\lambda\int_t^{t'}(u + A_{-i})\}$, the probability of no discovery by t', $1 - p + pQ_{t'}$, equals $(1-p)/(1-p_{t'})$. Dividing payoffs by $1-p$ removes the stochastic part of the problem. The first term in (6.1) is the player's own expected prize net of the safe payoff. The second is the value e_N of the known resource, which every player receives at a discovery regardless of who makes it. Player i's effort enters the second term in the same way as anybody else's, since every player benefits from a discovery. The reduced game is a deterministic differential game in which each player's payoff is linear in own effort.

For a continuously differentiable candidate value $W(x,t)$ define the *private gain*

$$G(x,t) = r\,m(x,t) + \lambda\big(e_N e^x - W_x(x,t)\big). \tag{6.2}$$

The gain is the marginal value of own effort. In the original units, with $v = s + (1-p)W$, the quantity $\phi = (1-p)G/r = (gp\rho - s) + (\lambda/r)p\big[s + e_N - v - (1-p)v_p\big]$ is the flow net return of the risky activity, plus the value of a discovery, minus the value lost by lowering the posterior. Against opponents who use a, player i's Hamilton–Jacobi–Bellman (HJB) equation reads

$$rW = W_t + \lambda K a\big(e_N e^x - W_x\big) + \max_{b\in[0,1]} b\,G, \tag{6.3}$$

so a best response uses full effort where $G > 0$ and waits where $G < 0$. In a symmetric equilibrium,

$$rW = W_t + a\big(NG - K\,r\,m\big), \qquad a \in \arg\max_{b\in[0,1]} bG. \tag{6.4}$$

To describe how the gain changes along paths of constant effort, define

$$\begin{aligned} \Theta_0(t) &= \lambda e_N + r\big(g\rho(t) - s\big) - g\eta\big(1-\rho(t)\big), & S_0(x,t) &= rs - \Theta_0(t)\,e^x, \\ \Theta_1(t) &= \Theta_0(t) + \lambda K\big(g\rho(t) - s\big), & S_1(x,t) &= rs - \Theta_1(t)\,e^x. \end{aligned} \tag{6.5}$$

Lemma 6.2 (Transport of the gain)**.** *Suppose that, on an open set, all players use the constant effort $a \in \{0,1\}$ and W is twice continuously differentiable and satisfies (6.4). Then along the path of that effort, $\dot{x} = -\lambda N a$, $\dot{t} = 1$,*

$$\frac{d}{dt}G = r\,G + r\,S_a. \tag{6.6}$$

The sources S_0 and S_1 depend only on the state, so integrating (6.6) backward from a point at which the gain is known determines its sign along entire paths. Optimality forces the gain to vanish wherever effort is interior and, because the equilibrium value is continuously differentiable, also at the boundaries of regions of interior effort. In the regions of waiting and of full effort, the sign of the gain is then the sign of a discounted integral of $-S_0$ or $-S_1$, and the equilibrium can be built region by region from these integrals.

6.2 A verification theorem

A strategy that is nondecreasing in the belief, together with a continuously differentiable value that solves the symmetric HJB equation wherever the strategy is continuous, forms an equilibrium under the Filippov convention if the strategy jumps only on the myopic curve, where the gain vanishes, and prescribes there an equal share of the effort that keeps the state on the curve.

Theorem 4 (Verification)**.** *Let $a : \mathbb{R}\times[0,\infty) \to [0,1]$ be Borel and nondecreasing in x, and let W be continuous on $\mathbb{R}\times[0,\infty)$, continuously differentiable on $\mathbb{R}\times(0,\infty)$, with $|W(x,t)| \le C(1+e^x)$. Suppose that*

(i) at every (x,t) with $t>0$ at which $a(\cdot,t)$ is continuous, W satisfies (6.4);

(ii) at every (x,t) at which $a(\cdot,t)$ is discontinuous, $m(x,t)=0$ and $G(x,t)=0$;

(iii) at each such discontinuity the prescribed action is

$$a(x,t)=\frac{\eta[1-\rho(t)]}{\lambda N[\rho(t)-q]}.$$

Then the symmetric Markov strategy that uses $a(\operatorname{logit} p, t(\rho))$ for $0<p<1$, the known-state equilibrium at $p=1$ and zero effort at $p=0$ is a symmetric Markov perfect equilibrium under the Filippov convention from every initial state with $\rho<1$, and its value is $v=s+(1-p)W$. On the edge $\rho=1$ the same conclusion holds for a Borel strategy $a_\infty:\mathbb{R}\to[0,1]$ that is nondecreasing and a value $W_\infty\in C^1(\mathbb{R})$ with $|W_\infty(x)|\le C(1+e^x)$ that satisfy the stationary versions of (i)–(iii), with $W_t=0$ and zero prescribed effort at any discontinuity; there the value is $v(p,1)=s+(1-p)W_\infty(\operatorname{logit} p)$.

Hypotheses (ii)–(iii) place discontinuities on the myopic curve and prescribe the equal effort shares needed for a profile path. Since $m=G=0$ there, the verification inequality is independent of how effective effort is split (Lemma B.1). The HJB inequality also covers deviations that leave the curve.

6.3 The frontiers

High quality, $\rho>\rho_B$. Define the lower frontier

$$L(\rho)=\frac{rs}{\lambda\big[e_N-\eta h+(r+\eta)h\rho\big]}, \tag{6.7}$$

which decreases from $L(\rho_B)=q/\rho_B$ to the individual abandonment belief

$$\ell_N=L(1)=\frac{rs}{\lambda(rh+e_N)}. \tag{6.8}$$

$L(\rho)$ is where the waiting source S_0 vanishes. Let $y_m(\rho)=\operatorname{logit}(q/\rho)-\operatorname{logit} L(\rho)>0$ be the log-odds distance between the myopic curve and L, and let $y_H(\rho)\in(0,y_m(\rho))$ be the unique solution of

$$r\big(e^y-1-y\big)=\lambda K\big(1-e^{y-y_m(\rho)}\big). \tag{6.9}$$

The high-quality band is $L(\rho)<p<p_H(\rho)$ with $\operatorname{logit} p_H(\rho)=\operatorname{logit} L(\rho)+y_H(\rho)$, and in it each player uses

$$\alpha(p,\rho)=\frac{r\big(e^y-1-y\big)}{\lambda K\big(1-e^{y-y_m(\rho)}\big)},\qquad y=\operatorname{logit} p-\operatorname{logit} L(\rho). \tag{6.10}$$

At $\rho=1$ the game is an exponential-bandit game of the kind studied by Keller et al. (2005), in which the first success pays its finder h and gives every player the continuation surplus e_N. The band then corresponds to their symmetric equilibrium, and interior effort rises from zero at ℓ_N to one at the upper full-quality cutoff $\bar p_N=p_H(1)$.

Low quality, $E_N < \rho < \rho_*$. Let $\mathcal{E}$ be the curve formed by the sliding arc and the upper frontier p_H of the high-quality band. Consider a state whose full-effort path passes above the onset point $(\rho_*, q/\rho_*)$ of the sliding arc. In the coordinates of Section 6.1, this is a state on a line $x + \lambda N t = \xi$ with $\xi > \xi_* = x_m(t_*) + \lambda N t_*$, where $t_* = t(\rho_*)$. Full effort by all players brings such a state to $\mathcal{E}$. Let $\tau_{\mathcal{E}}(p, \rho)$ be the time it takes to do so, and define the *full-effort gain*

$$\Gamma(p, \rho) = -rs(1 - e^{-r\tau_{\mathcal{E}}}) + r\,\frac{p}{1-p}\int_0^{\tau_{\mathcal{E}}} e^{-\Lambda_N \sigma}\,\Theta_1(t(\rho) + \sigma)\,\mathrm{d}\sigma. \tag{6.11}$$

This is the private gain at (p, ρ) if all players use full effort until the state reaches $\mathcal{E}$. Under Condition 1 below, for each $\rho \in (E_N, \rho_*)$ there is a unique $p_H(\rho) \in (q/\rho, 1)$ at which Γ vanishes, and $\Gamma > 0$ above it (Lemma C.6). With $x_{Hc} = \operatorname{logit} p_H(\rho)$ and $x = \operatorname{logit} p$, define

$$\mathcal{N}(p, \rho) = rs\,(x - x_{Hc}) - \Theta_0(t(\rho))(e^x - e^{x_{Hc}}) + \lambda K\, m(x_{Hc}, t(\rho)). \tag{6.12}$$

The low-quality band is $p_L(\rho) < p < p_H(\rho)$, where $p_L(\rho) \in (q/\rho, p_H(\rho))$ is the unique zero of $\mathcal{N}(\cdot, \rho)$ there (Lemma C.7), and in it each player uses the effort $\mathcal{N}(p, \rho)/[\lambda K\, m(\operatorname{logit} p, t(\rho))]$.

To complete the description, set $p_L = p_H = q/\rho$ on $[\rho_*, \rho_B]$ and $p_L = L$ on $(\rho_B, 1]$. Let $\rho_c \in [E_N, \rho_*)$ be the quality at which the full-effort path through the onset point $(\rho_*, q/\rho_*)$, traced backward in time, meets the myopic curve again, and $\rho_c = E_N$ if it does not do so at a quality above E_N. Let $\rho_\Theta = (\eta\rho_B + rq)/(\eta + r)$.

Condition 1. The primitives satisfy:

(R1) regular onset of sliding: $(r + \lambda K + \eta)A_m(\rho_*) > Nr$;

(R2) slow upper frontier at high quality: for every $\rho \in (\rho_B, 1)$,

$$\Psi(\rho) = r(e^{y_H(\rho)} - 1)\Big[\lambda N - \frac{\eta(1-\rho)}{\rho - \rho_\Theta}\Big] + \lambda^2 K\, e^{y_H(\rho) - y_m(\rho)}(N - A_m(\rho)) > 0;$$

(R3) no early activation below the myopic curve: $\Gamma(q/\rho, \rho) < 0$ for every $\rho \in (E_N, \rho_c)$.

Section D gives a closed form for the gain in (R3) (Lemma D.1), proves Proposition 2, and verifies the baseline with exact rational bounds (Lemma D.3).

6.4 The construction region by region

In the low-quality band, along both frontiers, the probability $1 - p$ that the resource is unproductive grows linearly in $\rho - E_N$ (Lemma C.8 gives the slopes). As quality rises, the full-effort region expands downward. Its boundary is the upper frontier p_H of the low-quality band, where the full-effort gain Γ vanishes, and above it full effort until the state reaches the curve $\mathcal{E}$ is a best response. In contrast to the upper frontier of the high-quality band, this frontier falls faster than full effort lowers the posterior, so it is never reached from above and instead overtakes the states in the band below it.

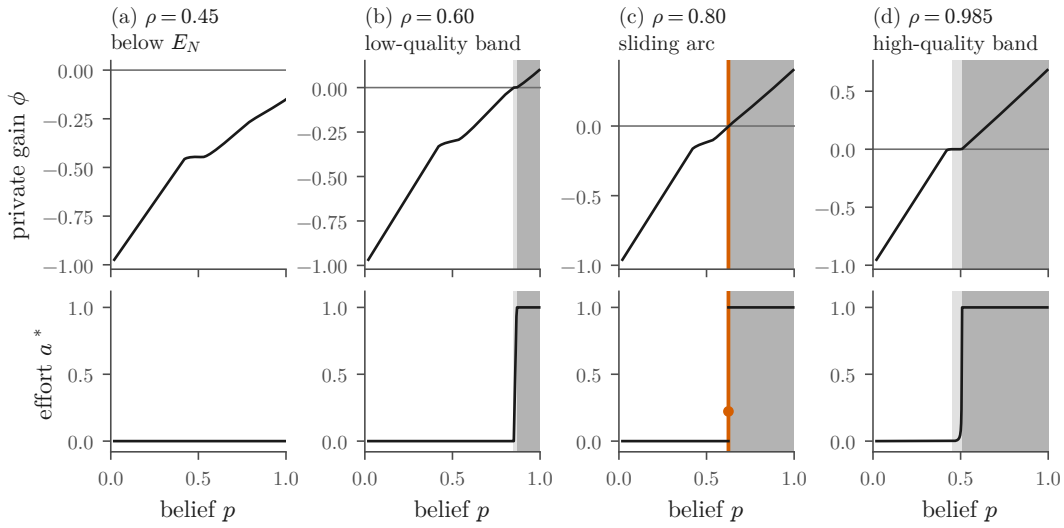


Figure 11: Each player's effort follows the sign of her private gain from effort. Top: the private gain $\phi = (1-p)G/r$, a player's net marginal value of her own effort, on four slices of quality. Bottom: each player's equilibrium effort a^* on the same slices. Players wait where the gain is negative (white), use full effort where it is positive (grey), and use interior effort where it is zero (light grey bands). (a) Below E_N the gain is negative at every belief. (b) The low-quality band. (c) At qualities on the sliding arc's range, the gain changes sign at $p = q/\rho$, where effort jumps, and on the arc itself each player uses A_m/N (dot). (d) The high-quality band. Baseline parameters, as in Figure 1.

Inside the band, effort is interior and rises from zero at p_L to one at p_H. Indifference ($G = 0$) pins down interior effort, and continuity with the full-effort region at p_H determines the value in the band. At the baseline parameters the band is at most about 0.17 wide in log odds, and it shrinks to zero at ρ_* (Figure 5b).

On the sliding arc the gain is zero, and the value along the arc satisfies the linear equation $rW = \dot{W} + \lambda e_N e^{x_m} A_m$, whose only flow term is the expected value of a discovery.

At ρ_B the waiting source S_0 vanishes on the myopic curve. Above ρ_B, waiting just below the curve is no longer optimal, because there the value of the information that experimentation produces exceeds its flow cost. A band of interior effort opens below the curve, with lower frontier L where $S_0 = 0$. In this band each player is indifferent and effort rises with the belief. As $\rho \to 1$ the band converges to the band at full quality, and along a path in the band the posterior converges to ℓ_N.

Figure 11 shows the private gain on four slices of quality, with the signs described in Theorem 2(c), and the effort that each sign implies. On slice (c) the gain changes sign exactly at $p = q/\rho$, where the slice meets the sliding arc.

The proof in Section C builds the value backward from high quality through the sliding arc and the low-quality band, and Figure 13 shows the geometry. Conditions (R1)–(R3) govern the frontier locations and slopes. Table 2 reports the baseline margins, and Figure 14 shows where the conditions pass numerical tests across parameters.

7 Trajectories and sliding

At intermediate qualities, every regular symmetric equilibrium switches from waiting to full effort on the myopic curve, and against such a switch some unilateral deviations have no ordinary trajectory.

7.1 Trajectories

Because Markov strategies can be discontinuous, we must specify how the state moves when players use them, both on the equilibrium path and after a unilateral deviation. For $0 < p < 1$ we use the log odds $x = \operatorname{logit} p = \log[p/(1-p)]$, in which (2.2) becomes $\dot{x} = -\lambda A$. Fix the opponents' strategies $(a_j)_{j \neq i}$ and let player i use an arbitrary measurable effort schedule $u : [0, \infty) \to [0, 1]$ until the next opportunity. With t measured from the initial date and $\rho(t)$ the solution of (2.1) from ρ_0, an ordinary trajectory solves

$$\dot{x}(t) = -\lambda\Big[u(t) + \textstyle\sum_{j \neq i} a_j\big(\operatorname{expit} x(t), \rho(t)\big)\Big], \tag{7.1}$$

where expit is the inverse of logit. When the opponents' strategies are discontinuous in p, (7.1) may have no absolutely continuous solution, and Section 7.2 shows that this problem is unavoidable for regular symmetric profiles at intermediate qualities. We therefore define trajectories in the sense of Filippov (1988), in which, at a discontinuity, the opponents' aggregate effort may take any value in the convex hull of its essential limits. For fixed t, let

$$\mathcal{A}_{-i}(x, t) = \bigcap_{\delta > 0} \bigcap_{\mu(Z) = 0} \overline{\operatorname{co}}\Big\{\textstyle\sum_{j \neq i} a_j(\operatorname{expit} y, \rho(t)) : \ |y - x| < \delta, \ y \notin Z\Big\}, \tag{7.2}$$

where μ is Lebesgue measure. A *trajectory* is an absolutely continuous function $x(\cdot)$ such that, for almost every t,

$$\dot{x}(t) = -\lambda\big[u(t) + A_{-i}(t)\big] \quad \text{for some } A_{-i}(t) \in \mathcal{A}_{-i}(x(t), t). \tag{7.3}$$

We call $A_{-i}(t) = -\dot{x}(t)/\lambda - u(t)$ the opponents' *effective* aggregate effort. Where the opponents' strategies are continuous in p, the set $\mathcal{A}_{-i}$ is a singleton and a trajectory is an ordinary solution of (7.1). Where they jump, a trajectory may cross the discontinuity or, if the drift points toward it from both sides, *slide* along it with the effective effort that keeps the state on the discontinuity. The convention specifies how the opponents' effective effort responds to a deviation on a discontinuity curve. When the opponents' strategies are nondecreasing in p, as in the equilibrium we construct, every ordinary trajectory is a trajectory in this sense, and by Lemma 7.1 below it is the only one. For such strategies the convention matters only where ordinary trajectories fail to exist. Our equilibrium results refer to the game it defines.

Formally, a profile is well posed against unilateral deviations if, for every player, every initial state and every measurable u, (7.3) has a unique trajectory on $[0, \infty)$.

Lemma 7.1. *If every opponent's strategy is nondecreasing in p for each ρ, the profile is well posed against unilateral deviations. Two trajectories generated by schedules u and $\tilde{u}$ from the same initial*

state satisfy $|x(t) - \tilde{x}(t)| \le \lambda \int_0^t |u(\sigma) - \tilde{u}(\sigma)| \, \mathrm{d}\sigma$.

The proof in Section B uses Filippov's existence theorem and monotonicity for uniqueness and the stated bound.

Formally, a profile path of a symmetric strategy a is an absolutely continuous function $x(\cdot)$ with $\dot{x}(t) = -\lambda N a(\operatorname{expit} x(t), \rho(t))$ for almost every t. A profile path can slide along a discontinuity curve only if each player's prescribed action on the curve is the equal share of the aggregate effort that keeps the state on it. The equilibrium of Theorem 2 prescribes exactly this action on its sliding arc (Table 1), and the proof of Theorem 4 shows that its profile path exists, is unique, and is a trajectory of every player's unilateral system.

Given the no-discovery trajectory, the stochastic process is standard. Conditional on $\theta = 1$, opportunities of player i and of the opponents arrive with intensities $\lambda u(t)$ and $\lambda A_{-i}(t)$. An opportunity resets the state to $(1, 0)$, after which the edge $p = 1$ is invariant and quality follows (2.1) between subsequent opportunities. All intensities are bounded by λN, so the jump process is nonexplosive. At $p = 0$ the belief stays at zero and no opportunities occur. At $p = 1$ it stays at one. Player i's payoff depends on the opponents' effective efforts only through their aggregate, because the continuation after a discovery does not depend on who makes it.

7.2 Necessity of sliding

With ρ_B and ρ_* as in (4.1), on the myopic curve $S_0 = s\eta(\rho_B - \rho)/(\rho - q)$, so waiting just below the curve has a positive source exactly when $\rho < \rho_B$. Similarly, $S_1 = s\eta(\rho_B - \rho)/(\rho - q) - \lambda K s$ on the curve, so full effort just above it has a negative source exactly when $\rho > \rho_*$. Keeping the state on the myopic curve requires the aggregate effort $A_m(\rho)$ in (4.2), which is below N for every $\rho \ge \rho_*$. Full effort therefore lowers the posterior faster than the myopic curve falls, and waiting lets the curve catch up with the state.

With t again the quality clock of Section 6.1, we call a symmetric profile (a, W) *regular* on a strip $\mathbb{R} \times (t_1, t_2)$ if W is continuously differentiable there; a is continuous outside finitely many C^1 curves $x = \zeta_k(t)$ and has one-sided limits in x on them; (6.4) holds wherever a is continuous; and, for each t, effort is zero at sufficiently low beliefs and full at sufficiently high beliefs.

Proposition 5 (Necessity of sliding)**.** *Let $\rho(t_1)$ and $\rho(t_2)$ lie in $[\rho_*, \rho_B]$, and let (a, W) be regular on $\mathbb{R} \times (t_1, t_2)$. Then, except possibly at points of the curves ζ_k, $a(x, t) = 0$ for $x < x_m(t)$ and $a(x, t) = 1$ for $x > x_m(t)$. Consequently:*

(a) under the profile, every Filippov trajectory that meets the myopic curve stays on it, with aggregate effort $A_m(\rho)$, as long as $\rho < \rho(t_2)$;

(b) if trajectories are required to be absolutely continuous solutions of (7.1), then, whatever actions the profile assigns on the myopic curve and on the curves ζ_k, there is a state and a measurable deviation for which no trajectory exists.

In particular, on this range of qualities no regular symmetric profile is a Markov perfect equilibrium with ordinary trajectories.

The first part of the proposition rests on two exclusion arguments. A band of interior effort that starts below the myopic curve must start where $S_0 \leq 0$, which happens below the curve only if $\rho > \rho_B$. A band that ends at full effort above the curve must end where $S_1 \geq 0$, which happens above the curve only if $\rho < \rho_*$. An action discontinuity anywhere else would contradict differentiability of the value. Subtracting (6.4) on the two sides of a jump gives $NG = Krm$, and optimality on both sides gives $G = 0$, so $m = 0$ at any jump, and only the switch at the myopic curve remains. Part (b) follows because, when the drift points toward a curve from both sides, an absolutely continuous solution started on the curve cannot leave it, so its velocity must equal that of the curve, and a small deviation in own effort makes this impossible.

Remark 2 (Uniqueness). For qualities in $[\rho_*, \rho_B]$, Proposition 5 implies that every symmetric equilibrium in the regular class defined above waits below the myopic curve and uses full effort above it, away from its interfaces, as the equilibrium of Theorem 2 does. On the curve itself, a profile path exists only if each player's action is A_m/N. Outside this range of qualities we do not know whether the equilibrium is unique. The equilibrium of Theorem 2 continues the sliding arc into each adjacent range of qualities with a single band of interior effort.

7.3 Implementation by frequent updates

Remark 3 (Interpretation of sliding). On the sliding arc the prescribed effort jumps from waiting to full effort, and effective effort keeps the state on the curve. One way to implement sliding is to let players revise their actions only at common dates and keep them fixed between revisions. The state then chatters around the curve, and its average motion is the sliding motion (Figure 6b). For the equilibrium constructed here, Proposition 6 makes this approximation precise. Sliding occurs only on the myopic curve, where the flow payoff from own effort equals s, so the way effective effort is split among the players does not affect their flow payoffs (Lemma B.1).

Between update dates the opponents' efforts are constant, so every deviation generates an ordinary trajectory. A deviating player may change effort at any time, including between update dates.

Proposition 6 (Sampled implementation)**.** *Assume Condition 1, let a^* be the equilibrium of Theorem 2, and let $\Delta > 0$. Consider the profile in which, before a discovery, all players update their effort at the dates $k\Delta$ after the initial date, $k = 0, 1, \ldots$, to the effort prescribed by a^* at the current state and hold it until the next update date, and in which they use the known-state equilibrium after a discovery. For every compact set $\mathcal{C} \subset (0,1) \times [0,1)$ of initial states there is $\varepsilon_\Delta(\mathcal{C})$, with $\varepsilon_\Delta(\mathcal{C}) \to 0$ as $\Delta \downarrow 0$, such that from every initial state in $\mathcal{C}$ no adapted unilateral deviation raises a player's payoff by more than $\varepsilon_\Delta(\mathcal{C})$. As $\Delta \downarrow 0$, the no-discovery path of the profile converges to the profile path of a^*, uniformly on bounded time intervals, and its payoffs converge to those of a^*, uniformly*

on $\mathcal{C}$; against a fixed measurable unilateral schedule, the no-discovery path converges uniformly on bounded time intervals to the trajectory defined by (7.3).

The proof is in Section B.5.

8 Hidden failures

When failed opportunities are hidden, competition can dissipate a player's whole surplus from a discovery, and players then learn less than when every opportunity is public.

Suppose that actions and the opportunities that pay, which we call *successes*, remain public, but that the opportunities that pay nothing, which we call *failures*, are observed by nobody, not even the player who finds them. Beliefs therefore stay common. A success shows that the resource is productive and resets quality to zero, so it sends the state to $(1, 0)$, as every opportunity does in Section 2, and in this section a discovery is the first success. Between successes quality is hidden, because nobody sees when a failure resets it. In the fishing example, guides see the fish that are landed but not the encounters that disturb a site without a catch. Let ψ denote the *expected quality* of the resource, conditional on its being productive and on the public history. Initial quality is known, so $\psi_0 = \rho_0$. Between successes, aggregate effort moves the belief and expected quality deterministically.

Lemma 8.1 (Filter)**.** *Fix a state (p, ψ) with $p > 0$, and suppose that aggregate effort follows a measurable path A_t until the next success. Then, until the next success,*

$$\dot{p}_t = -\lambda A_t \psi_t p_t (1 - p_t), \qquad \dot{\psi}_t = (1 - \psi_t)(\eta - \lambda A_t \psi_t) \tag{8.1}$$

for almost every t, and the next success arrives with intensity $\lambda A_t p_t \psi_t$ and sends the state to $(p, \psi) = (1, 0)$.

The belief falls more slowly than in (2.2), because only a success shows that the resource is productive. Effort now lowers expected quality, for two reasons. A failure resets quality unseen, and the absence of a success suggests that quality is low. Recovery raises expected quality at the rate $\eta(1 - \psi)$. Expected quality summarizes everything the players need to know about quality, because the motion (8.1) of the state, the intensity of a success and expected payoffs all depend on the conditional distribution of quality only through its mean. A Markov strategy is therefore a Borel function of (p, ψ), and Definition 1 applies with ψ in place of ρ. Section H defines trajectories.

Working the resource pays more than the safe activity if and only if $p\psi > q$, so the myopic curve of Section 2 becomes $p\psi = q$. The aggregate effort that keeps the state on the curve is

$$\hat{A}_m(\psi) = \frac{\eta(1 - \psi)}{\lambda\psi(1 - q)}. \tag{8.2}$$

On the edge $p = 1$ the curve is the point $\psi = q$, where the aggregate effort $\hat{A}_m(q) = \eta/(\lambda q)$ holds

expected quality still. The other $N-1$ players can supply this effort by themselves if and only if $\lambda(N-1)q \geq \eta$, that is $(N-1)s \geq \eta h$. In that case players wait where working the resource pays less than the safe activity, use full effort where it pays more, and slide along the whole curve on which it pays the same.

Proposition 7 (Hidden failures)**.** *Suppose that $(N-1)s \geq \eta h$. The symmetric Markov strategy that waits where $p\psi < q$, uses full effort where $p\psi > q$ and uses $\hat{A}_m(\psi)/N$ where $p\psi = q$ is a symmetric Markov perfect equilibrium under the Filippov convention, and it has the following properties.*

(a) After every success, expected quality recovers to q and is held there until the next success, and each player's value is s whenever $\psi \leq q$. In particular, a discovery is worth nothing to an individual player.

(b) Before a discovery, each player's value is s wherever $p\psi \leq q$, and the state slides along the whole myopic curve.

(c) From every initial state with $0 < p_0 < 1$, along the no-discovery path the belief converges to $\min\{p_0, q\}$, and a discovery occurs with probability $(p_0 - q)_+/(1-q)$. Neither depends on the initial quality.

If instead $(N-1)s < \eta h < Ns$, the strategy is not an equilibrium.

Hiding failures changes the game after a discovery. When failures are observed, effort lowers quality only through opportunities, each of which resets it to zero for everybody to see. Players can then only choose when to harvest the recovering resource, and the race to harvest first leaves each of them the surplus $e_N > 0$ (Proposition 1). When failures are hidden, effort lowers expected quality gradually, so players can hold it at q, where a unit of effort earns exactly the safe flow. Above q, a player who holds back gives up her share of a positive flow. If the other $N-1$ players can hold expected quality at q by themselves, her holding back does not stop expected quality from falling to q. Full effort is then a best response above q, and competition after a discovery dissipates each player's whole surplus, as in the open-access fishery of Gordon (1954). When $(N-1)s < \eta h < Ns$, a player who holds back just above q lets expected quality rise, which is worth more to her than her share of the flow, so the strategy fails, and when $\eta h > Ns$ even full effort by all players cannot hold expected quality at q. We do not know the equilibrium when $(N-1)s < \eta h$.

Since a discovery is worth nothing to a player, information has no private value before a discovery. Above the curve, a player who holds back saves little of the surplus for later, because the other players' full effort pushes the state to the curve anyway, and when $(N-1)s \geq \eta h$ this saving is worth less than her current surplus. Each player's equilibrium effort therefore maximizes her current flow payoff, as if she ignored the future, and we call the equilibrium *myopic*. When failures are observed, the equilibrium approaches this form only as the number of players grows (Proposition 4), because $e_N > 0$ for every N.

When $(N-1)s \geq \eta h$, hiding failures therefore reduces learning (Figure 12). When failures are observed, learning stops at $\ell_N < q$ (Theorem 3), because a discovery is worth e_N to a player. When

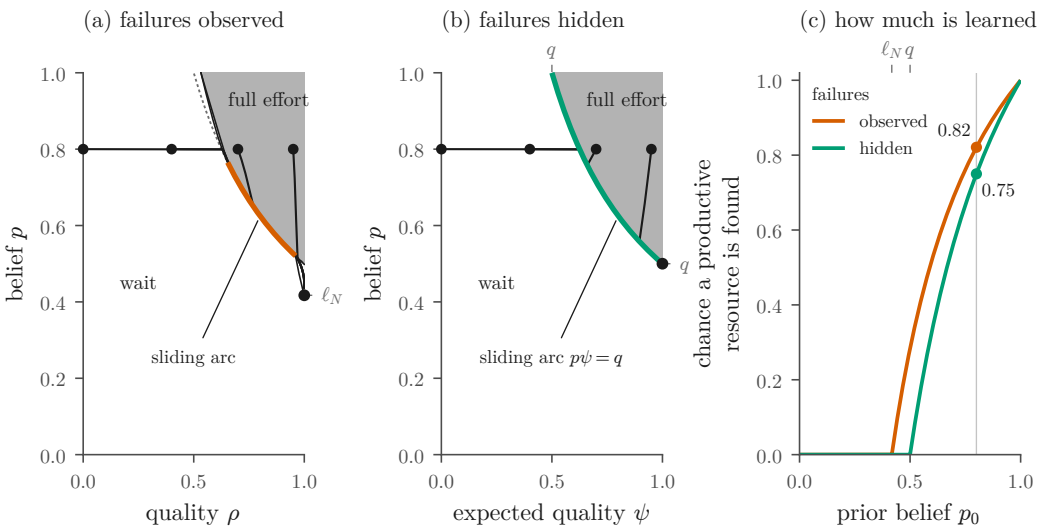


Figure 12: Hiding failures stops learning at a higher belief, from every initial quality. (a) The equilibrium of Theorem 2, in which failures are observed, and its no-discovery paths (black) from the belief $p_0 = 0.8$ and the qualities 0, 0.4, 0.7 and 0.95 (dots on the line $p = 0.8$). Shading shows individual effort, from white (wait) to grey (full effort). Thin black curves are the frontiers of the bands of interior effort, as in Figure 5, the vermilion curve is the sliding arc, and the dotted curve continues the myopic curve $p = q/\rho$. Every path approaches the dot at the belief $\ell_N = 0.42$ and full quality. (b) The equilibrium of Proposition 7, in which failures are hidden, and its no-discovery paths from the same states, with expected quality ψ on the horizontal axis. Players wait below the myopic curve $p\psi = q$, use full effort above it and slide along all of it, so the whole curve is the sliding arc (green). Full effort lowers expected quality, so paths above the curve move to the left. Every path approaches the dot at the belief $q = 0.5$ and full expected quality. (c) The probability that the resource is found if it is productive, which is the probability of a discovery divided by p_0, as a function of the prior belief p_0 and from any initial quality, when failures are observed (vermilion) and when they are hidden (green). The vertical line marks the prior 0.8 of panels (a) and (b). Baseline parameters, as in Figure 1, for which $(N-1)s = \eta h$.

they are hidden, learning stops at q, the belief at which a player who ignores discoveries stops. From priors above q, the belief approaches q along the myopic curve, because the aggregate effort $\hat{A}_m(\psi)$ that keeps the state on the curve falls as expected quality recovers and vanishes at full expected quality. At the baseline parameters, Condition 1 holds and $(N-1)s = \eta h$, so both equilibria exist. From a prior of 0.8, a discovery occurs with probability 0.66 when failures are observed and 0.60 when they are hidden, so a resource that is productive is found with probability 0.82 and 0.75. From priors between $\ell_N = 0.42$ and $q = 0.5$, players never experiment when failures are hidden.

The proof in Section H follows Sections 6 and 7. Without a success, the odds $p/(1-p)$ and $p\psi/(1-p)$ follow a linear system, and their difference decays at the rate η whatever the players do, so its logarithm serves as the quality clock. Above the myopic curve, the private gain satisfies a transport equation like (6.6), whose source is negative when $(N-1)s \geq \eta h$. On the edge $p = 1$, where every success returns expected quality to zero, a bound that applies afresh after each success completes the proof.

9 Conclusion

For experimentation on a resource that the experiments themselves deplete, we solved the planner's problem and, under three conditions on primitives and with Filippov trajectories, constructed a symmetric Markov perfect equilibrium. Players exploit too early when productivity is likely and quality is moderate. From uncertain priors above the planner's abandonment belief, they also learn too little overall. Initial quality affects only when they experiment before a discovery. In the fishing example, guides return before a promising site's catchability has recovered and stop exploring uncertain sites that a planner would keep exploring. At intermediate qualities the state slides along the myopic curve, dissipating the flow rent. Frequent updates of actions give ordinary trajectories and approximate Nash equilibria (Proposition 6).

The three conditions of our existence theorem hold in large groups but can fail when recovery is fast or players are very impatient. The frontiers of our construction can then detach from the sliding arc, cross the myopic curve or fall faster than full effort lowers the posterior, and an equilibrium would need a different structure. For all primitives, symmetric Nash equilibria in deterministic schedules exist, and every equilibrium in deterministic schedules reaches the same terminal belief, $\min\{p_0, \ell_N\}$, as the Markov perfect equilibrium (Section F). It is open whether a Markov perfect equilibrium exists outside our conditions or with ordinary trajectories, whether it is unique outside the intermediate range of qualities or the regular class, what asymmetric equilibria, which in exponential-bandit games can improve on the symmetric one (Keller et al., 2005), look like in this model, and at what rate the error of the sampled implementation vanishes. Partial resets of quality or inconclusive news, which the model rules out, would change both the learning problem and the game after a discovery.

What players observe also shapes how much they learn. When failures are hidden and $(N-1)s \geq \eta h$, there is an equilibrium in which competition after a discovery dissipates the whole rent of the known resource and learning stops at q, above ℓ_N (Section 8). We do not know the equilibria when $(N-1)s < \eta h$, or when each player privately observes her own failures, so that beliefs are no longer common.

A Proofs for Section 3

A.1 The known-state equilibrium (Proposition 1)

On $\{p = 1\}$ write $u_0 = s + e_N$, $D(\rho) = \eta(1-\rho)$ and $b(\rho) = g\rho - s$. Against the symmetric threshold profile, the coefficient of own effort in a player's normalized Bellman equation is

$$\phi(\rho) = b(\rho) + \frac{\lambda}{r}\big[u_0 - u(\rho)\big].$$

For the symmetric profile with effort a, the candidate value satisfies $Du' = r(u-s) - rab + \lambda N a(u-u_0)$ almost everywhere. Waiting below a candidate threshold and full effort above it give the two branches in (3.3); on the active branch, the stated solution is the only one that remains bounded as $\rho \to 1$. Matching $u(E_N-) = u(E_N+)$ and private indifference $\phi(E_N) = 0$ eliminate e_N and give (3.1); substitution gives (3.2).

Existence and uniqueness of the root. Let $F(\varrho)$ denote the right side of (3.1) with ϱ in place of E_N. With $z = 1 - \varrho$, $H = 1 - z^\beta$ and $c_0 = \lambda\eta/L_N$,

$$F'(\varrho) = 1 + c_0\Big\{\frac{H}{r + \lambda K H} - \frac{r\beta z^\beta}{(r + \lambda K H)^2}\Big\} \geq 1 - \frac{\lambda}{L_N} > 0.$$

Since $F(0) = 0$, $F(1) = 1$ and $F(\varrho) < \varrho$ for $0 < \varrho < 1$, the equation $F(E_N) = q$ has a unique root in $(q, 1)$.

Optimality. On the active branch $\phi' = g - (\lambda/r)u' = g(1 - \lambda/L_N) > 0$. On the waiting branch u' increases with quality, and $u'(E_N-) = (rg/L_N)\, r/(r + \lambda K H_N) < rg/L_N = u'(E_N+)$, so $\phi' > g(1 - \lambda/L_N)$ there as well. Hence $\phi < 0$ below the threshold and $\phi > 0$ above it. The value equation then implies $\mathcal{L}^{\hat a} u + r[s + \hat a\,(g\rho - s)] - ru \leq 0$ almost everywhere for every own effort $\hat a \in [0, 1]$, where $\mathcal{L}^{\hat a} u = \eta(1-\rho)u' + \lambda[\hat a + Ka(\rho)][u(0) - u(\rho)]$. Between opportunities, quality follows $\dot\rho = \eta(1-\rho)$ independently of effort and crosses E_N at positive speed, so the action assigned at E_N is irrelevant. Integrating the expected jump increments against the bounded predictable intensity and applying the inequality gives, for every adapted deviation and every T,

$$u(\rho) \geq \mathbb{E}\Big[\int_0^T re^{-rt}\big\{s + \hat a_t(g\rho_t - s)\big\}\,\mathrm{d}t + e^{-rT}u(\rho_T)\Big],$$

with equality under the profile. Boundedness of u removes the terminal term as $T \to \infty$, and the arrival rate is bounded by λN, so opportunities do not accumulate.

Comparison with the planner. The planner's threshold equation (3.4) subtracts a strictly larger quantity than (3.1) at every interior candidate threshold, and both right sides are strictly increasing; hence $E_N < R_N$. Let $U(\rho) = V(1, \rho)$ and let $G_P(\rho) = (g\rho - s)/N + (\lambda/r)[U(0) - U(\rho)]$ be the planner's effort coefficient (Section A.2). The equilibrium uses aggregate effort N on (E_N, R_N), where $G_P < 0$; the residual of the maximized planner equation under this effort is $-rNG_P > 0$. Hence $U(0) - u(0) = D_N - e_N$ equals the expected discounted integral of this residual along the

equilibrium path from $\rho = 0$. A compact subinterval of (E_N, R_N) is reached before a reset with positive probability and occupied for positive time, so the integral is positive and $e_N < D_N$. □

A.2 The planner's known-state problem

Throughout the rest of this appendix, $\kappa = \lambda N$ and $\Omega = g(r + \eta)$. The known-state planner value $U(\rho) = V(1, \rho)$ is given by (3.3) with (E_N, e_N) replaced by (R_N, D_N) from (3.4).

Uniqueness of the root. Let $F_N(\varrho)$ be the right side of (3.4) with ϱ in place of R_N, and $z = 1 - \varrho$. Then $F_N'(\varrho) = 1 + \frac{\kappa\eta}{r(r+\kappa+\eta)}[1 - (1+\beta)z^\beta] \geq \frac{r+\eta}{r+\kappa+\eta} > 0$, $F_N(0) = 0$, $F_N(1) = 1$ and $F_N(q) < q$, so the root is unique in $(q, 1)$. The two branches of U and their derivatives agree at R_N, where both one-sided derivatives equal $rg/(r + \kappa + \eta)$.

Optimality. The known-state Bellman equation is

$$rU(\rho) = \eta(1-\rho)U'(\rho) + rs + \max_{0 \leq A \leq N} A\Big[\frac{r}{N}(g\rho - s) + \lambda\{U(0) - U(\rho)\}\Big].$$

The coefficient of aggregate effort divided by r is $G_P(\rho) = (g\rho - s)/N + (\lambda/r)[U(0) - U(\rho)]$. Substitution gives $G_P(R_N) = 0$. On the active branch $G_P' = \Omega/[N(r + \kappa + \eta)] > 0$. On the waiting branch $U'' = \beta(\beta+1)D_N(1-\rho)^{-\beta-2} > 0$ and U' increases to $rg/(r+\kappa+\eta)$, so $G_P' \geq \Omega/[N(r+\kappa+\eta)]$ there too. Hence waiting maximizes the Bellman equation below R_N and full effort above it. The bounded value U satisfies the equation under the threshold policy, and the same verification argument as in Section A.1 proves optimality against every adapted aggregate effort process.

An identity used below follows from (3.4):

$$\Pi_N = \Omega(1 - R_N)\frac{r + \kappa}{r + \kappa + \eta} > 0. \tag{A.1}$$

A.3 A concave problem in the survival path

Before the first discovery quality follows $\rho_t = 1 - (1 - \rho)e^{-\eta t}$. For a measurable schedule $A : [0, \infty) \to [0, N]$ define $Q_t = \exp\{-\lambda \int_0^t A_u \, \mathrm{d}u\}$, the probability of no opportunity by t conditional on productivity. The unconditional survival probability is $1 - p + pQ_t$, and the first-discovery density is $p\lambda A_t Q_t$. For $p > 0$ put

$$c = \frac{rs(1-p)}{p}, \qquad B = \Omega(1 - \rho), \qquad \Pi(t) = \Pi_N - Be^{-\eta t}.$$

Lemma A.1 (Survival representation)**.** *For $0 < p \leq 1$,*

$$V(p, \rho) - s = \frac{pr}{N\lambda} \sup_Q \int_0^\infty e^{-rt}\big[\Pi(t)(1 - Q_t) + c \log Q_t\big] \, \mathrm{d}t, \tag{A.2}$$

where the supremum is over locally absolutely continuous Q with $Q_0 = 1$ and $-\kappa Q \leq Q' \leq 0$ almost everywhere. For $0 < p < 1$ the objective is strictly concave in Q.

Proof. After a discovery the planner obtains $s + D_N$. Let $J(A)$ be the payoff of schedule A until discovery followed by the optimal known-state continuation. The first-discovery decomposition gives

$$J(A) - s = \frac{p}{N} \int_0^\infty e^{-rt} A_t \big[(\kappa D_N + rg\rho_t - rs)Q_t - c\big] \, \mathrm{d}t.$$

Put $k(t) = \kappa D_N + rg\rho_t - rs$; then $k' - rk = -r\Pi$. Substituting $A_t Q_t = -Q_t'/\lambda$ and $A_t = -(\log Q_t)'/\lambda$ and integrating by parts gives (A.2): the bounds $e^{-\kappa t} \le Q_t \le 1$ and $|\log Q_t| \le \kappa t$ remove the terms at infinity, and the terms at zero produce the constant in $\Pi(1 - Q)$. Conversely every feasible Q induces the measurable effort $-Q'/(\lambda Q) \in [0, N]$. Before discovery, conditional on any randomization independent of the hidden state, a policy determines one no-opportunity schedule, so randomization cannot improve on the supremum. The feasible set is convex, and for $c > 0$ the second derivative of the integrand in Q is $-c/Q^2 < 0$; distinct continuous paths differ on an interval, so the objective is strictly concave. □

A.4 The frontier and the optimal paths

Define

$$B_c = \frac{\kappa}{\kappa + \eta} \Pi_N, \quad c_c = \frac{\eta}{\kappa + \eta} \Pi_N, \quad \rho_c^P = 1 - \frac{B_c}{\Omega}, \quad p_c^P = \frac{rs}{rs + c_c}. \tag{A.3}$$

By (A.1), $R_N < \rho_c^P < 1$. For $a > 0$ and $T > 0$ let

$$I_a(T) = \frac{1 - e^{-(r+a)T}}{r + a} - e^{-aT} \frac{1 - e^{-rT}}{r}, \qquad H(T) = \frac{I_\kappa(T)}{I_{\kappa+\eta}(T)}, \qquad f(T) = e^{-\kappa T} \big[1 - H(T) e^{-\eta T}\big]. \tag{A.4}$$

The low-quality branch of the frontier is

$$\rho_b(T) = 1 - \frac{\Pi_N}{\Omega} H(T), \qquad p_b(T) = \frac{rs}{rs + \Pi_N f(T)}, \qquad 0 < T < \infty, \tag{A.5}$$

and the high-quality branch is (3.5) on $[\rho_c^P, 1]$.

These formulas define a continuous strictly decreasing graph $p = p^P(\rho)$ on $[R_N, 1]$. Writing $e^{-at} - e^{-aT} = \int_t^T a e^{-au} \, \mathrm{d}u$ gives $I_a(T) = \frac{a}{r} \int_0^T e^{-au}(1 - e^{-ru}) \, \mathrm{d}u$. Let $w(u) = e^{-(\kappa+\eta)u}(1 - e^{-ru})$ and $M_P(T) = \int_0^T e^{\eta u} w \, \mathrm{d}u / \int_0^T w \, \mathrm{d}u$, so that $H = \frac{\kappa}{\kappa+\eta} M_P$. Then $M_P'(T) = w(T)[e^{\eta T} - M_P(T)] / \int_0^T w > 0$, and $H(T) e^{-\eta T} < \kappa/(\kappa + \eta)$. The expansions $I_a(T) = aT^2/2 + o(T^2)$ and $I_a(\infty) = 1/(r + a)$ give $H(0+) = \kappa/(\kappa + \eta)$ and $H(\infty) = (r + \kappa + \eta)/(r + \kappa)$. Also $f'(T) = e^{-\kappa T}\{-\kappa + (\kappa + \eta) H e^{-\eta T} - H' e^{-\eta T}\} < 0$, $f(0+) = \eta/(\kappa + \eta)$ and $f(\infty) = 0$. Hence $\rho_b' < 0 < p_b'$; (A.1) gives the endpoint $(R_N, 1)$, and $T \to 0$ gives (ρ_c^P, p_c^P). On the high branch $dp^P/d\rho = -rs\Omega/[rs + \Pi_N - \Omega(1 - \rho)]^2 < 0$; it joins the low branch at ρ_c^P and reaches $(1, p_N^P)$.

Construction. At $p = 0$ choose $A = 0$; at $p = 1$ use the known-state policy. For $0 < p \le p_N^P$ choose $Q \equiv 1$. For $p_N^P < p < 1$, define B_b by $B_b = \Pi_N - c$ if $c \ge c_c$ and by $B_b = \Pi_N H(T_b)$ with $\Pi_N f(T_b) = c$ if $c < c_c$. Wait for

$$t_w = \max\Big\{0, \frac{1}{\eta} \log \frac{B}{B_b}\Big\}$$

($t_w = 0$ if $B = 0$). With $B_w = Be^{-\eta t_w}$ and time $u = t - t_w$, use full effort for the duration T given by the largest solution $T \geq 0$ of

$$c = e^{-\kappa T}\big[\Pi_N - B_w e^{-\eta T}\big], \tag{A.6}$$

and then

$$\widehat{Q}_u = \frac{c}{\Pi_N - B_w e^{-\eta u}}, \qquad u \geq T, \tag{A.7}$$

which corresponds to aggregate effort $A^s(\rho_t)$ in (3.5).

Proposition 8 (Global verification)**.** *The path just constructed is the unique maximizer of* (A.2) *for* $0 < p < 1$*, and together with the known-state policy it implements an optimal policy from every state.*

Proof. If $c \geq \Pi_N$, the derivative of the integrand is $c/Q - \Pi(t) \geq 0$ for $Q \leq 1$, so $Q \equiv 1$ is optimal. Let $0 < c < \Pi_N$.

Feasibility. For $0 \leq \xi \leq B_b$ let $h_\xi(u) = e^{-\kappa u}(\Pi_N - \xi e^{-\eta u})$. Its derivative has the sign of $-\kappa\Pi_N + (\kappa + \eta)\xi e^{-\eta u}$, which is nonincreasing, so h_ξ has at most one interior maximum and tends to zero. At a low-branch start, $h_{B_b}(T_b) = c$ and $h'_{B_b}(T_b) < 0$ because $H(T_b)e^{-\eta T_b} < \kappa/(\kappa + \eta)$; if $\xi < B_b$, then $h_\xi(T_b) > c$, so there is a later root on the decreasing branch. At a high-branch start, $h_{B_b}(0) = c$ and $h'_{B_b}(0) \leq 0$, and $h_\xi(0) > c$ for $\xi < B_b$. Since $B_w \leq B_b$, the duration T exists. At the terminal root $\eta B_w e^{-\eta T}/(\Pi_N - B_w e^{-\eta T}) \leq \kappa$, and the left side decreases afterward, so $-\kappa\widehat{Q} \leq \widehat{Q}' \leq 0$ in the final phase.

Multiplier. Let A^P be the constructed aggregate effort, so that λA^P is the hazard of an opportunity, and let $Q_t^P = \exp(-\lambda \int_0^t A^P)$ be its survival path. We construct an absolutely continuous z_0 with

$$z_0' = e^{-rt}(\Pi - c/Q^P) + \lambda A^P z_0,$$

with $z_0 \geq 0$ while waiting, $z_0 \leq 0$ under full effort, $z_0 = 0$ in the final phase, and vanishing terminal products. Concavity and integration by parts then bound the gain of any feasible Q, up to a positive factor, by

$$\int_0^\infty e^{-rt}(c/Q^P - \Pi)(Q - Q^P)\,\mathrm{d}t = \int_0^\infty z_0(t)\big[Q_t' + \lambda A_t^P Q_t\big]\,\mathrm{d}t \leq 0,$$

because $Q' \leq 0$ while waiting and $Q' + \kappa Q \geq 0$ under full effort. In time u after waiting, set $\widehat{z} = 0$ in the final phase and, under full effort,

$$\widehat{z}(u) = -e^{\kappa u}\int_u^T e^{-(r+\kappa)v} d(v)\,\mathrm{d}v, \qquad d(v) = \Pi_N - B_w e^{-\eta v} - ce^{\kappa v}.$$

Then $\widehat{z}(T) = 0$ and the multiplier equation holds. Since $d'' < 0$ and the terminal root lies on the decreasing branch, d is nonnegative before T or changes sign once from negative to positive. At a low-branch start, substitution of (A.4)–(A.5) gives $\int_0^T e^{-(r+\kappa)v} d(v)\,\mathrm{d}v = 0$; at a high-branch start $T = 0$. For $B_w < B_b$, let $\mathcal{I}(\xi)$ be this integral at the root selected for coefficient ξ; then $\mathcal{I}'(\xi) = -\int_0^{T(\xi)} e^{-(r+\kappa+\eta)v}\,\mathrm{d}v < 0$, so $\mathcal{I}(B_w) \geq 0$. The sign pattern of d then gives a nonnegative integral

from every $u \le T$, so $\widehat{z} \le 0$. Set $z_0(t) = e^{-rt_w}\widehat{z}(t-t_w)$ after waiting and $z_0(t) = \int_t^{t_w} e^{-rv}[c-\Pi(v)]\,\mathrm{d}v$ during waiting. If $t_w > 0$ then $B_w = B_b$, so $\widehat{z}(0) = -\mathcal{I}(B_b) = 0$ and z_0 is continuous at t_w. In addition, $\Pi(t_w) - c = d(0)$, which is zero at a high-branch start and negative at a low-branch start, because there d changes sign once, from negative to positive, and $\mathcal{I}(B_b) = 0$; since Π is increasing, $z_0 \ge 0$ while waiting. Terminal products vanish because $0 \le Q \le 1$ and $z_0 = 0$ after $t_w + T$. Strict concavity gives uniqueness.

Feedback form. By uniqueness, the no-discovery tail at any date is the optimal path from the current state, with $c_t = c/Q_t^P$ and $B_t = Be^{-\eta t}$. The waiting phase ends when $B_t = B_b(c_t)$, that is, exactly when the state reaches the frontier: for $c_t \ge c_c$ this is $rs(1-p_t)/p_t = \Pi_N - \Omega(1-\rho_t)$, the high branch, and for $c_t < c_c$ it is the point $(\rho_b(T_b), p_b(T_b))$ of the low branch. A state strictly above the frontier, or on its low branch, has $t_w = 0$ and $T > 0$ (on the low branch $T = T_b$), so the planner uses full effort; a state on the high branch has $t_w = T = 0$, and the final phase keeps $c_t = \Pi_N - \Omega(1-\rho_t)$, that is, $p_t = p^P(\rho_t)$. This is the stationary policy in Theorem 1. At $p = 1$ optimality follows from Section A.2; at $p = 0$ the safe action is optimal; every opportunity leads to $(1, 0)$, whose continuation is optimal, so the first-discovery decomposition proves optimality in the original problem. □

A.5 Learning and the size of the group

For $p_N^P < p < 1$ the optimal path satisfies $Q_\infty = c/\Pi_N$, so total effort is $\lambda^{-1}\log(\Pi_N/c) = \lambda^{-1}[\operatorname{logit} p - \operatorname{logit} p_N^P]$, the limiting posterior is p_N^P, and the probability of a discovery is $p(1-Q_\infty) = (p - p_N^P)/(1 - p_N^P)$. This proves (3.6). In addition, $V(p,\rho) = s$ if and only if $p \le p_N^P$: for $p > p_N^P$, choose t_0 with $\Pi(t_0) > c$ and use a short burst of effort at t_0; the derivative $c/Q - \Pi$ of the integrand is negative along the altered tail, so the objective rises.

Lemma A.2 (Planner and group size)**.** *R_N and ND_N increase strictly with N and p_N^P decreases strictly. Let $R_\infty \in (q,1)$ solve $q = R_\infty - \frac{\eta}{r}(1-R_\infty)[1-(1-R_\infty)^\beta]$. Then $R_N \to R_\infty$, $ND_N \to C_P = \eta h(1-R_\infty)^{1+\beta} > 0$, and $p_N^P \to rs/[\lambda(C_P + rh)] < q$.*

Proof. The coefficient $\lambda N\eta/(rL_N)$ in (3.4) increases strictly with N and the right side is increasing in the threshold, so R_N rises; the limiting equation has a unique root in $(q,1)$ because its right side has derivative $(1+1/\beta)[1-(1-t)^\beta] > 0$. Substitution in (3.4) gives the limit of ND_N. For monotonicity of ND_N, write V_N for the planner's value with N players and note that $N[V_N(1,0) - s] = \sup_{0 \le A \le N} \mathbb{E}\int_0^\infty re^{-rt}A_t(g\rho_t - s)\,\mathrm{d}t$ is the value of a problem whose feasible set expands with N; strictness follows by raising effort slightly on an interval of qualities above R_N that is reached with positive probability, where the marginal gain of the smaller planner is strictly positive. Finally $p_N^P = rs/[\lambda(ND_N + rh)]$ decreases strictly and converges to $rs/[\lambda(C_P + rh)] < rs/(\lambda rh) = q$. □

B Foundations: proofs for Sections 6 and 7

B.1 Well-posedness (Lemma 7.1)

Fix t. Because each $a_j(\operatorname{expit} y, \rho(t))$ is nondecreasing in y, so is the opponents' aggregate $\Upsilon_t(y) = \sum_{j\neq i} a_j(\operatorname{expit} y, \rho(t))$; its one-sided limits $\Upsilon_t(x\pm)$ exist, and the set (7.2) is the interval $\mathcal{A}_{-i}(x,t) = [\Upsilon_t(x-), \Upsilon_t(x+)]$. The right side of the inclusion (7.3) is the Filippov regularization of a function that is Borel in (t,x) and bounded by λN; the endpoints $\Upsilon_t(x\pm)$ are monotone limits of Borel functions, so the set-valued map is measurable in t and upper semicontinuous with compact convex values in x. Hence a trajectory exists on every bounded interval (Filippov, 1988, Ch. 2), and boundedness extends it to $[0,\infty)$.

Let x and $\tilde{x}$ be trajectories for schedules u and $\tilde{u}$ from the same initial state, with opponents' effective efforts A_{-i} and $\tilde{A}_{-i}$. If $x(t) > \tilde{x}(t)$, then every element of $\mathcal{A}_{-i}(x(t),t)$ is at least $\Upsilon_t(x(t)-) \geq \Upsilon_t(\tilde{x}(t)+)$, which is at least every element of $\mathcal{A}_{-i}(\tilde{x}(t),t)$; hence $(x-\tilde{x})(A_{-i}-\tilde{A}_{-i}) \geq 0$ almost everywhere. Therefore

$$\frac{d}{dt}(x-\tilde{x})^2 = -2\lambda(x-\tilde{x})\big[(u-\tilde{u}) + (A_{-i}-\tilde{A}_{-i})\big] \leq 2\lambda|x-\tilde{x}|\,|u-\tilde{u}|.$$

Since $|x-\tilde{x}|$ is absolutely continuous and its derivative vanishes almost everywhere on $\{x=\tilde{x}\}$, this gives $\frac{d}{dt}|x-\tilde{x}| \leq \lambda|u-\tilde{u}|$ almost everywhere, so $|x-\tilde{x}|(t) \leq \lambda\int_0^t |u-\tilde{u}|$. With $u=\tilde{u}$ this is uniqueness. □

B.2 The reduction (Lemma 6.1)

Measure calendar time $\sigma \geq 0$ from the initial state, let (x_σ) be the no-discovery trajectory, $A_\sigma = u_\sigma + A_{-i,\sigma}$, $Q_\sigma = \exp(-\lambda\int_0^\sigma A)$ and $S_\sigma = 1-p+pQ_\sigma$ the probability of no opportunity by σ. Conditional on no opportunity, the posterior is $p_\sigma = pQ_\sigma/S_\sigma$ and the intensity of the first opportunity is $\lambda p_\sigma A_\sigma$. At the first opportunity every player's continuation value under the known-state equilibrium is $s+e_N$; a deviation after it cannot raise player i's continuation (Proposition 1). Splitting the payoff at the first opportunity,

$$J^i = \int_0^\infty e^{-r\sigma}S_\sigma\Big[r\{s+u_\sigma(gp_\sigma\rho_\sigma - s)\} + \lambda p_\sigma A_\sigma(s+e_N)\Big]\,\mathrm{d}\sigma.$$

Since $\frac{d}{d\sigma}(e^{-r\sigma}S_\sigma) = -e^{-r\sigma}S_\sigma(r+\lambda p_\sigma A_\sigma)$ and $e^{-r\sigma}S_\sigma \to 0$, the terms proportional to s integrate to s, so

$$J^i - s = \int_0^\infty e^{-r\sigma}S_\sigma\big[ru_\sigma(gp_\sigma\rho_\sigma - s) + \lambda p_\sigma A_\sigma e_N\big]\,\mathrm{d}\sigma.$$

Finally $S_\sigma(1-p_\sigma) = S_\sigma - pQ_\sigma = 1-p$, so $S_\sigma = (1-p)/(1-p_\sigma)$. Substituting, and using $(gp\rho-s)/(1-p) = m$ and $p/(1-p) = e^x$, gives (6.1) with $t' = t+\sigma$. The same computation applies at $\rho = 1$, where quality stays at one until the first opportunity, with $e^x(g-s)-s$ in place of $m(x,t')$. □

B.3 Transport of the gain (Lemma 6.2)

For constant effort $a \in \{0,1\}$, put $\mathcal{D}_a = \partial_t - \lambda N a \partial_x$. Equation (6.4) reads $rW = \mathcal{D}_a W + a(rm + \lambda N e_N e^x)$, so differentiation in x gives $\mathcal{D}_a W_x = rW_x - a(rm_x + \lambda N e_N e^x)$. Hence

$$\mathcal{D}_a G = rm_t - \lambda K a r m_x - r\lambda W_x = rG + r\big[m_t - \lambda K a m_x - \lambda e_N e^x - rm\big] = rG + rS_a,$$

because $m_t - \lambda e_N e^x - rm = S_0$ and $S_1 = S_0 - \lambda K m_x$. □

B.4 The verification theorem (Theorem 4)

Lemma B.1 (Payoffs on the sliding arc)**.** *Suppose a trajectory spends a set $\mathcal{T}$ of times on a curve on which $m = 0$. Then, for almost every $t \in \mathcal{T}$, player i's reduced flow payoff is $-e_N e^{x(t)} \dot{x}(t)$, whatever the split of effective effort between the deviator and the opponents.*

Proof. The flow in (6.1) is $ru\,m + \lambda e_N e^x (u + A_{-i})$; on the curve the first term vanishes and $\lambda(u + A_{-i}) = -\dot{x}$ almost everywhere. □

Proof of Theorem 4. Since $a(\cdot, t)$ is nondecreasing, Lemma 7.1 gives well-posedness against unilateral deviations. It remains to show that the profile path exists and is unique, and that no deviation is profitable.

Existence of the profile path. Write $\mathcal{B}(x,t) = [a(x-,t), a(x+,t)]$. The existence and monotonicity proof of Lemma 7.1, with the aggregate coefficient N, gives a unique solution of

$$\dot{x}(t) \in -\lambda N \mathcal{B}(x(t), t).$$

At a continuity point of $a(\cdot,t)$ this is $\dot{x} = -\lambda N a(x,t)$. At a discontinuity, condition (ii) puts the state on the smooth myopic curve. An absolutely continuous function has derivative zero almost everywhere on its zero set; applying this to $x(t) - x_m(t)$ shows that, at almost every time spent at a discontinuity,

$$\dot{x}(t) = \dot{x}_m(t) = -\lambda A_m(\rho(t)) = -\lambda N a(x(t), t),$$

where the last equality is hypothesis (iii). Thus this path follows the prescribed pointwise actions almost everywhere, so it is a profile path. Every profile path solves the inclusion, since $a(x,t) \in \mathcal{B}(x,t)$, so the profile path is unique. Along it, $u(t) = a(x(t),t)$ and $A_{-i}(t) = K a(x(t),t)$ solve every player's unilateral inclusion, because a monotone function's pointwise value lies between its one-sided limits. At full quality the same argument uses a stationary myopic point and $A_m(1) = 0$.

Optimality against unilateral controls. Fix an initial state (x_0, t_0), a deviation u, its trajectory $x(\cdot)$ and the opponents' effective effort A_{-i}. The function $\Xi(t) = e^{-r(t-t_0)} W(x(t), t)$ is absolutely continuous on compact subintervals of (t_0, ∞) and continuous at t_0, and almost everywhere

$$\Xi'(t) + e^{-r(t-t_0)}\big[ru\,m + \lambda e_N e^x (u + A_{-i})\big] = e^{-r(t-t_0)}\,\iota(t), \qquad \iota = W_t - rW + \lambda A_{-i}(e_N e^x - W_x) + uG.$$

Continuity points. If $a(\cdot,t)$ is continuous at $x(t)$, then $\mathcal{A}_{-i} = \{Ka\}$ and $A_{-i} = Ka$. Since $a(NG - Krm) = aG + \lambda Ka(e_N e^x - W_x)$, (6.4) gives $\iota = (u-a)G \le 0$, because a maximizes bG over $b \in [0,1]$.

Discontinuity points. If $a(\cdot,t)$ is discontinuous at $x^* = x(t)$, then by (ii) $m = G = 0$, so $e_N e^x - W_x = (G - rm)/\lambda = 0$ and $\iota = W_t - rW$. The continuity points of the monotone function $a(\cdot,t)$ are dense; along a sequence of them converging to x^*, (6.4) gives $rW - W_t = a(NG - Krm) \to 0$, because a is bounded and $NG - Krm$ is continuous and vanishes at x^*. Hence $\iota = 0$.

Thus $\iota \le 0$ almost everywhere, with equality if $u = a$ wherever a is continuous. Integrating first from $t_0 + \varepsilon$ to T and then sending $\varepsilon \downarrow 0$ gives, by continuity of W and boundedness of the flow on this path,

$$W(x_0,t_0) \ge \int_{t_0}^{T} e^{-r(t-t_0)}\left[ru\,m + \lambda e_N e^x (u + A_{-i})\right] \mathrm{d}t + e^{-r(T-t_0)} W(x(T),T).$$

Because x is nonincreasing, $|W(x(T),T)| \le C(1+e^{x_0})$, and the last term vanishes as $T \to \infty$. Hence the reduced payoff of any deviation schedule is at most $W(x_0,t_0)$, with equality for the prescribed profile path established above. By Lemma 6.1, applied conditional on the deviator's independent randomization and then in expectation, $J^i \le s + (1-p)W = v$, with equality in equilibrium. At $p = 1$ the known-state equilibrium is optimal against all deviations (Proposition 1); at $p = 0$ there are no opportunities and $gp\rho - s < 0$, so zero effort is optimal. At $\rho = 1$ the state stays at $\rho = 1$ until the first opportunity, and the same argument applies with W_∞ in place of W and $W_t \equiv 0$. □

B.5 Implementation by frequent updates (Proposition 6)

Proof of Proposition 6. Fix a compact set $\mathcal{C}$ of initial states, measure elapsed time by $\sigma \ge 0$, and write the quality clock as $t_0 + \sigma$. Let

$$\mathcal{B}(x,t) = [a^*(x-,t), a^*(x+,t)]$$

be the completed graph of the equilibrium feedback. It has compact convex values and is upper semicontinuous jointly in (x,t): off the closed sliding arc a^* is jointly continuous (Theorem 2(b)), so $\mathcal{B}$ is single-valued and continuous there, and on the closed arc, including its endpoints, $\mathcal{B} = [0,1]$. This joint property is the regularity used here in addition to the hypotheses of Lemma 7.1.

Paths. Against a measurable deviation u_Δ, let b_Δ be the opponents' sampled aggregate effort. The no-discovery path is the ordinary solution of

$$\dot{x}_\Delta = -\lambda(u_\Delta + b_\Delta), \qquad b_\Delta(\sigma) = Ka^*\big(x_\Delta(k\Delta), t_0 + k\Delta\big) \quad \text{for } k\Delta \le \sigma < (k+1)\Delta.$$

These paths are Lipschitz with constant λN. Along any sequence $\Delta \downarrow 0$, a subsequence of x_Δ converges uniformly on bounded intervals, and u_Δ and b_Δ converge weakly star in L^∞, to limits x, u and b; passing to the limit in the integral equation gives $\dot{x} = -\lambda(u+b)$. The sampled

state differs from the current state by at most $\lambda N\Delta$, and the sampled time by at most Δ. By upper semicontinuity, for every σ the value $b_\Delta(\sigma)$ lies between $Ka^*(x(\sigma)-, t_0+\sigma) - d_\Delta(\sigma)$ and $Ka^*(x(\sigma)+, t_0+\sigma) + d_\Delta(\sigma)$, where $d_\Delta(\sigma) \to 0$ for every σ and $|d_\Delta| \le N$. Integrating these bounds against nonnegative test functions in L^1 and using dominated convergence gives

$$b(\sigma) \in K\mathcal{B}(x(\sigma), t_0+\sigma) \quad \text{for almost every } \sigma,$$

so the limit is a trajectory (7.3) of the unilateral system with schedule u. For a fixed schedule u, uniqueness (Lemma 7.1) identifies every subsequential limit, so the whole family converges.

Payoffs. Lemma 6.1 applies to sampled profiles, because its proof uses only the no-discovery path. Since $\lambda(u+b)e^x = -\frac{d}{d\sigma}e^x$ and $e^{-r\sigma}e^{x(\sigma)} \to 0$, integrating the second term of (6.1) by parts gives

$$W^i = e_N e^{x_0} + r\int_0^\infty e^{-r\sigma}\left[u(\sigma)m(x(\sigma), t_0+\sigma) - e_N e^{x(\sigma)}\right] \mathrm{d}\sigma. \tag{B.1}$$

The right side depends on the opponents only through the path. On bounded intervals, uniform convergence of x_Δ and weak-star convergence of u_Δ imply convergence of the integral, even when the schedules vary with Δ; and because $x_\Delta(\sigma) \le x_0$, the integrand is bounded and the tails beyond any date $\bar\sigma$ are bounded by a constant times $e^{-r\bar\sigma}$, uniformly over initial states in $\mathcal{C}$. Hence payoffs converge along every convergent subsequence.

Equilibrium payoffs and incentives. When all players use the sampled feedback, the same argument applied to the full profile gives limits that solve $\dot x \in -\lambda N\mathcal{B}(x,t)$. This inclusion has a unique solution, the profile path of Theorem 4, so the sampled no-discovery path converges to it and the sampled equilibrium payoff converges to v^*. The compactness argument applies verbatim when the initial states $(p_n, \rho_n) \in \mathcal{C}$ vary with $\Delta_n \downarrow 0$ and converge to some (p,ρ): the limit inclusion has a unique solution from (p,ρ), so the sampled equilibrium payoffs converge to $v^*(p,\rho)$. Together with continuity of v^* (Theorem 2(c)), this gives uniform convergence on $\mathcal{C}$. Suppose that the largest gain from a deviation does not converge to zero uniformly on $\mathcal{C}$. Then there are $\Delta_n \downarrow 0$, initial states in $\mathcal{C}$ converging to some $(p,\rho) \in \mathcal{C}$, and deviations whose gains are bounded below by some $\varepsilon > 0$. Before the first opportunity, and conditional on the deviator's independent randomization, each deviation is a measurable schedule; after it, the known-state equilibrium bounds every deviation. Passing to a convergent subsequence, the argument above yields a unilateral trajectory from (p,ρ) with payoff at least $v^*(p,\rho)+\varepsilon$, which contradicts Theorem 4. □

B.6 Necessity of sliding (Proposition 5)

Lemma B.2 (Order of thresholds)**.** *$q < E_N < \rho_m < \rho_* < \rho_B < 1$ and $1-\rho_B < \lambda(1-q)/L_N$, where $\rho_m = (\eta+\lambda Nq)/(\eta+\lambda N)$ is the quality at which $A_m(\rho_m) = N$. The functions Θ_0 and Θ_1 are strictly increasing, and $\Theta_0(t_E) < \Theta_1(t_E) < 0 < \Theta_1(t_*)$, where t_E and t_* are the quality clocks of E_N and ρ_*, and*

$$\int_0^\infty e^{-\Lambda_N u}\,\Theta_1(t+u)\,\mathrm{d}u = \phi^1(\rho(t)) \quad \textit{for all } t, \tag{B.2}$$

where $\phi^1(\rho) = g(1-\lambda/L_N)(\rho - E_N)$ is the private gain ϕ on the active branch of the known-state equilibrium; in particular the integral vanishes at $t = t_E$.

Proof. The inequality $q < E_N$ is part of Proposition 1. By (3.1), since $H_N/(r+\lambda K H_N) < 1/(\lambda K)$ and $z_N < 1-q$, $E_N - q < \eta(1-q)/(KL_N) < \eta(1-q)/(\eta+\lambda N) = \rho_m - q$. From (3.2), $e_N \le \eta g z_N^{1+\beta}/L_N < \eta g(1-q)/L_N$. Hence $1-\rho_B = e_N/(\eta h) < \lambda(1-q)/L_N < \lambda(1-q)/(\eta+\lambda N)$, which gives both $\rho_B > q$ and, after rearrangement, $(\eta+\lambda N)\rho_B > \eta + \lambda K + \lambda q$, which is equivalent to $\rho_* > \rho_m$. Since ρ_* is a strict convex combination of q and ρ_B, $q < \rho_* < \rho_B$; $\rho_B < 1$ because $e_N > 0$.

Both $\Theta_0(t) = C_0 - \Omega e^{-\eta t}$ and $\Theta_1(t) = C_1 - C_2 e^{-\eta t}$, with $C_0 = \lambda e_N + r(g-s)$, $\Omega = g(r+\eta)$, $C_1 = \lambda e_N + (r+\lambda K)(g-s)$ and $C_2 = g(r+\lambda K+\eta)$, are strictly increasing. The left side of (B.2) equals $C_1/\Lambda_N - C_2(1-\rho)/L_N$. On the active branch of (3.3), $\phi(\rho) = g\rho - s + (\lambda/r)[u_0 - u(\rho)]$ is affine with slope $g(1-\lambda/L_N)$ and vanishes at E_N. Expanding both affine functions of ρ and using $L_N = \Lambda_N + \eta$ and $(r+\lambda K) = \Lambda_N - \lambda$ shows that they coincide; this is (B.2). Write $I(t)$ for the left side. Then $I'(t) = \Lambda_N I(t) - \Theta_1(t)$, and at t_E, $I = 0$ and $I' = g(1-\lambda/L_N)\eta(1-E_N) > 0$, so $\Theta_1(t_E) < 0$; since $gE_N > s$, $\Theta_0(t_E) = \Theta_1(t_E) - \lambda K(gE_N - s) < \Theta_1(t_E)$. On the myopic curve $S_1 = s\eta(\rho_B - \rho)/(\rho - q) - \lambda K s$, which is strictly decreasing in ρ and vanishes at ρ_*. At t_*, $S_1 = 0$ on the myopic curve, so $\Theta_1(t_*)e^{x_m(t_*)} = rs > 0$. □

Lemma B.3 (Jumps lie on the myopic curve)**.** *Let (a, W) be regular on a strip, and suppose that at a point (x^*, t) of one of the curves ζ_k the one-sided limits $a(x^*\pm, t)$ differ. Then $G(x^*, t) = 0$ and $m(x^*, t) = 0$.*

Proof. Taking limits of (6.4) from both sides, with W, W_t and G continuous, gives $(a(x^*+,t) - a(x^*-,t))(NG - Krm) = 0$, so $NG = Krm$ at x^*. If $G(x^*,t) > 0$, then $G > 0$ nearby and optimality forces $a = 1$ on both sides; similarly if $G < 0$. Hence $G = 0$, and then $m = 0$. □

Proof of Proposition 5. We first show, slice by slice, that $a = 0$ below the myopic curve and $a = 1$ above it; parts (a) and (b) then follow. Fix $t \in (t_1, t_2)$, so that $\rho_* < \rho(t) < \rho_B$. Changing a at finitely many points of the slice preserves its one-sided traces and the HJB identities used below; the Filippov sets ignore these point values. Thus by Lemma B.3, after such a change $a(\cdot, t)$ is continuous on $(-\infty, x_m(t))$ and on $(x_m(t), \infty)$. Where $0 < a < 1$ on an open set, optimality gives $G = 0$, so $W_x = e_N e^x + rm/\lambda$ and $W = \Phi + \chi(t)$ with

$$\Phi(x,t) = e_N e^x + \frac{r}{\lambda}\big[e^x(g\rho(t) - s) - sx\big] \tag{B.3}$$

and χ continuously differentiable. Since $\Phi_t - r\Phi = (r/\lambda)(rsx - \Theta_0 e^x)$, (6.4) becomes

$$\lambda K\, a\, m = rs\, x - \Theta_0(t)e^x + \nu(t), \qquad \nu(t) = \frac{\lambda}{r}(\chi' - r\chi). \tag{B.4}$$

This identity also holds throughout an interval (x_a, x_b) of a fixed time slice on which $0 < a < 1$ after the pointwise modification, even if the interval crosses exceptional curves. Set $\chi = W - \Phi$

and remove the finitely many exceptional points from (x_a, x_b). Every remaining point has an open neighborhood on which $G = 0$, so χ and χ_t are locally constant in x on each remaining subinterval. Continuous differentiability of W makes χ and χ_t continuous across the removed points, so their constants agree throughout (x_a, x_b). Hence $\nu = (\lambda/r)(\chi_t - r\chi)$ is constant in x on (x_a, x_b), and (B.4) extends across the removed points and to the endpoints of the interval by continuity.

No interior effort below the curve. Suppose $a(x', t) > 0$ for some $x' < x_m(t)$, and let $x_0 = \sup\{x \le x' : a(x, t) = 0\}$, which is finite because effort vanishes at low beliefs. By continuity $a(x_0, t) = 0$ and $0 < a < 1$ on an interval $(x_0, x_0 + \varepsilon)$, on which (B.4) gives $\lambda K a m = rs(x - x_0) - \Theta_0(t)(e^x - e^{x_0})$. The left side is negative there because $m < 0$ below the curve, and the right side vanishes at x_0, so its derivative $S_0(x_0, t)$ is nonpositive. Then $\Theta_0(t) > 0$, so $S_0(\cdot, t)$ is decreasing, and $S_0(x_0, t) > S_0(x_m(t), t) = s\eta(\rho_B - \rho)/(\rho - q) > 0$, a contradiction. Hence $a = 0$ below the curve.

Full effort above the curve. Suppose $a(x', t) < 1$ for some $x' > x_m(t)$. Since effort is one at high beliefs and $a(\cdot, t)$ is continuous above the curve, $x_3 = \inf\{x \ge x' : a(x, t) = 1\}$ is finite, $a(x_3, t) = 1$ and $a < 1$ on $[x', x_3)$. With $x_4 = \max\{x_m(t), \sup\{x < x_3 : a(x, t) \in \{0, 1\}\}\} < x_3$, we have $0 < a < 1$ on (x_4, x_3). Fixing $\nu(t)$ by $a(x_3, t) = 1$ in (B.4) gives

$$\lambda K\, m(x, t)(1 - a(x, t)) = \int_x^{x_3} S_1(y, t)\, \mathrm{d}y, \qquad x_4 < x < x_3, \tag{B.5}$$

because both sides vanish at x_3 and the derivative of the difference is zero. The left side is positive, so $S_1(x_3, t) \ge 0$. But $\Theta_1(t) > \Theta_1(t_*) > 0$ for $t > t_*$ (Lemma B.2), so $S_1(\cdot, t)$ is decreasing and $S_1(x_3, t) < S_1(x_m(t), t) = s\eta(\rho_B - \rho)/(\rho - q) - \lambda K s < 0$, a contradiction. Hence $a = 1$ above the curve.

(a) With $D = x - x_m(t)$, the symmetric drift is $\dot{D} = \lambda A_m > 0$ below the curve and $\dot{D} = \lambda(A_m - N) < 0$ above it, since $A_m < N$ on $[\rho_*, \rho_B]$ (Lemma B.2). The Filippov set on the curve is $-\lambda[0, N]$, which contains $-\lambda A_m$. After the pointwise modification the profile is nondecreasing on each slice, and Filippov sets ignore sets of measure zero, so by Lemma 7.1 the unique trajectory slides.

(b) Fix $t_0 < t_0'$ in (t_1, t_2) and let $I = [t_0, t_0']$. Let n be the number of exceptional curves, and add the myopic curve to them as $\zeta_0 = x_m$. Put

$$\underline{u}(t) = \max\{0, A_m(t) - K\}, \qquad \overline{u}(t) = \min\{1, A_m(t)\}, \qquad \delta(t) = \frac{\overline{u}(t) - \underline{u}(t)}{n + 3},$$

where $A_m(t) = A_m(\rho(t))$. A deviation $u \in (\underline{u}, \overline{u})$ makes the myopic curve attracting from both sides. Since $0 < A_m < N$, the continuous function δ has a strictly positive minimum on I. Consider the $n + 2$ feasible candidate controls $u_j(t) = \underline{u}(t) + j\,\delta(t)$, $j = 1, \dots, n + 2$, and, for each curve, the control that would keep an ordinary trajectory on it,

$$f_k(t) = -\frac{\dot{\zeta}_k(t)}{\lambda} - Ka(\zeta_k(t), t), \qquad k = 0, \dots, n.$$

These functions are measurable. Each $f_k(t)$ lies within distance less than $\delta(t)/3$ of at most one

candidate, and there are more candidates than curves; let $u(t)$ be the candidate with the lowest index that satisfies $|u(t) - f_k(t)| \geq \delta(t)/3$ for every k. This is a measurable control with values in $[\underline{u} + \delta, \overline{u} - \delta]$; extend it measurably beyond I.

Suppose that an ordinary absolutely continuous trajectory starts at $x(t_0) = x_m(t_0)$. On the set of times at which $x = \zeta_k$, the derivative of $x - \zeta_k$ vanishes almost everywhere, so $\dot{x} = \dot{\zeta}_k$; the ordinary equation would then require $u = f_k$. By the choice of u, every such set of times has measure zero. At almost every other time, the first part of the proposition gives $a = 0$ below the myopic curve and $a = 1$ above it. Hence, with $D = x - x_m$,

$$\dot{D} = \lambda(A_m - u) \geq \lambda\delta > 0 \quad \text{where } D < 0, \qquad \dot{D} = \lambda(A_m - u - K) \leq -\lambda\delta < 0 \quad \text{where } D > 0.$$

The positive part D^+ is absolutely continuous with $\frac{d}{dt}D^+ = \mathbf{1}_{\{D>0\}}\dot{D} \leq 0$ almost everywhere and $D^+(t_0) = 0$, so $D^+ = 0$ on I; in the same way $D^- = 0$. The trajectory therefore stays on the myopic curve throughout I, which contradicts the fact that it spends a set of times of measure zero on it. Hence no ordinary trajectory exists for this measurable deviation, whatever actions are assigned on the exceptional curves. □

C Construction of the equilibrium: proof of Theorem 2

Throughout this appendix we work in the coordinates (x, t) of Section 6.1, write $z = e^x$ and $\Lambda = \Lambda_N$, and denote by $t_E < t_m < t_* < t_B$ and t_c the quality clocks of $E_N, \rho_m, \rho_*, \rho_B$ and ρ_c (Lemma B.2). Both $\Theta_0(t) = C_0 - \Omega e^{-\eta t}$ and $\Theta_1(t) = C_1 - C_2 e^{-\eta t}$ are strictly increasing (Lemma B.2), with $C_0 = \lambda e_N + r(g - s)$, $C_1 = C_0 + \lambda K(g - s)$, $\Omega = g(r + \eta)$ and $C_2 = g(r + \lambda K + \eta)$. On the myopic curve,

$$S_0(x_m(t), t) = \frac{s\eta(\rho_B - \rho)}{\rho - q}, \qquad S_1(x_m(t), t) = \frac{s\eta(\rho_B - \rho)}{\rho - q} - \lambda K s, \tag{C.1}$$

with $\rho = \rho(t)$; the first follows from $\Theta_0 + rs = \lambda[e_N - \eta h + (r + \eta)h\rho]$ and $e^{x_m} = q/(\rho - q)$, the second from $e^{x_m}(g\rho - s) = s$. The value is built in four steps: the high-quality region $t > t_B$, the sliding arc, the full-effort region, and the low-quality band; the waiting region is filled in last. Figure 13 shows the geometry.

C.1 High quality

For $t > t_B$ let $x_L(t) = \operatorname{logit} L(\rho(t))$, which is the zero of $S_0(\cdot, t)$: $\Theta_0(t)e^{x_L} = rs$. Let $y_m(t) = x_m(t) - x_L(t)$, $y_H(t)$ the solution of (6.9), and $x_H = x_L + y_H$. With $\rho_\Theta = (\eta\rho_B + rq)/(\eta + r)$, so that $\Theta_0(t) < 0$ if and only if $\rho(t) < \rho_\Theta$, one has $\rho_\Theta < \rho_B$ and, from $L(\rho) = rs/(\Theta_0 + rs)$,

$$\dot{x}_L(t) = -\frac{\eta(1 - \rho)}{\rho - \rho_\Theta}, \qquad e^{y_m} - 1 = \frac{\eta(\rho - \rho_B)}{r(\rho - q)}. \tag{C.2}$$

Lemma C.1 (High-quality band). *(a) $y_m(t) > 0$ for $t > t_B$, $y_m(t_B) = 0$, and (6.9) has a unique*

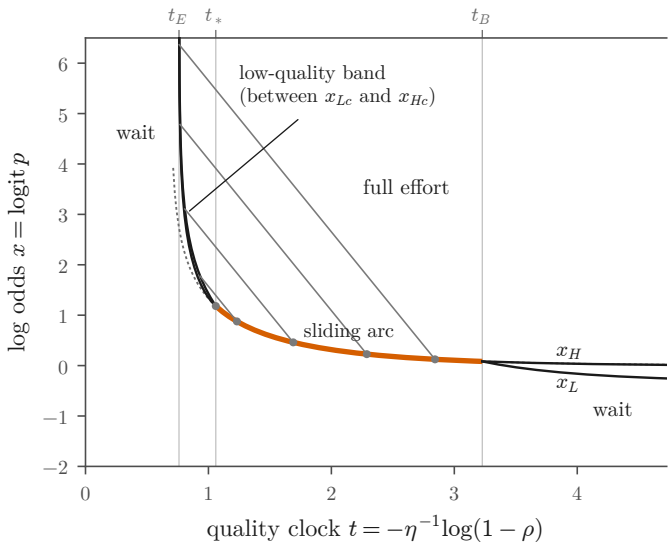


Figure 13: The construction in the coordinates (t, x), baseline parameters. The exit curve $\mathcal{E}$ consists of the sliding arc (thick) and the upper frontier x_H of the high-quality band. Straight lines are full-effort characteristics, traced backward from their exit points to the zeros of the full-effort gain, which form the upper frontier x_{Hc} of the low-quality band. The lower frontier x_{Lc} lies just below x_{Hc}.

root $y_H \in (0, y_m)$. The curves x_L and x_H are continuously differentiable and strictly decreasing, $x_L(t_B) = x_H(t_B) = x_m(t_B)$, and $x_L(t) \to x_\ell = \operatorname{logit} \ell_N$.

(b) $S_1(x_H(t), t) = -rs\, y_H(t) - \lambda K s < 0$.

(c) The band action (6.10) is continuous and strictly increasing in x on (x_L, x_H), with limits 0 at x_L and 1 at x_H.

(d) $\dot{x}_H + \lambda N$ has the sign of $\Psi(\rho(t))$ in (R2).

Proof. (a) By (C.2), $y_m > 0$ exactly when $\rho > \rho_B$. In (6.9) the left side increases from 0 and the right side decreases from $\lambda K(1 - e^{-y_m}) > 0$ to 0 on $[0, y_m]$, so the root is unique and interior, and it is a smooth function of y_m. $\dot{x}_L < 0$ by (C.2). Writing the condition $\alpha = 1$ as $F(x_H, t) = rs(e^y - 1 - y) - \lambda K(s - e^{x_H}(g\rho - s)) = 0$ with $y = x_H - x_L$, we have $F_x = rs(e^y - 1) + \lambda K e^{x_H}(g\rho - s) > 0$ and $F_t = -rs(e^y - 1)\dot{x}_L + \lambda K e^{x_H} g\eta(1 - \rho) > 0$, so $\dot{x}_H = -F_t/F_x < 0$. At $\rho = 1$, $x_L = \operatorname{logit} \ell_N$. (b) Using $\Theta_0 e^{x_L} = rs$ and the condition $\alpha = 1$, $\lambda K e^{x_H}(g\rho - s) = \lambda K s - rs(e^{y_H} - 1 - y_H)$, so $S_1(x_H) = rs - rse^{y_H} - \lambda K e^{x_H}(g\rho - s) = -rsy_H - \lambda K s$. (c) The numerator of (6.10) increases and the denominator decreases in y. (d) $\dot{x}_H + \lambda N = (\lambda N F_x - F_t)/F_x$, and $\lambda N F_x - F_t = s\,\Psi$ after substituting (C.2) and $e^{x_H}(g\rho - s) = se^{y_H - y_m}$, $\lambda N(g\rho - s) - g\eta(1 - \rho) = \lambda(g\rho - s)(N - A_m)$. □

Values. In the band, put $W = \Phi + \chi_T(t)$ with Φ from (B.3) and

$$\chi_T(t) = -\int_t^\infty e^{-r(v-t)} \varpi(v)\, dv, \qquad \varpi(v) = \frac{r^2 s}{\lambda}(1 - x_L(v)), \tag{C.3}$$

the unique bounded solution of $\chi' - r\chi = \varpi$; every other solution differs from it by a multiple of e^{rt} and would violate the growth bound of Theorem 4. Then $G = 0$ in the band and, by (B.4) with $\nu(t) = rs(1 - x_L(t))$, (6.4) holds with $a = \alpha$. At $t = \infty$ the same formulas with $x_L = x_\ell$ and $\chi_T = -(rs/\lambda)(1 - x_\ell)$ give the stationary band at $\rho = 1$; there $W(x_\ell) = 0$.

C.2 The sliding arc

On $[t_*, t_B]$ let W_{arc} solve

$$W'_{\mathrm{arc}}(t) = rW_{\mathrm{arc}}(t) - \lambda e_N e^{x_m(t)} A_m(t), \qquad W_{\mathrm{arc}}(t_B) = \Phi(x_m(t_B), t_B) + \chi_T(t_B), \tag{C.4}$$

where $\lambda e^{x_m} A_m = \eta q(1 - \rho)/(\rho - q)^2$. This is the value of sliding: on the arc each player's reduced flow is $\lambda e_N e^{x_m} A_m$ (Lemma B.1). We assign to the arc the gradient $W_x = e_N e^{x_m}$, so that $G = 0$ there; Lemma C.10 shows that this is the derivative of the value from both sides.

C.3 The exit curve and the full-effort gain

Let $\mathcal{E}$ be the graph of $x_{\mathcal{E}}(t) = x_m(t)$ for $t_* \le t \le t_B$ and $x_{\mathcal{E}}(t) = x_H(t)$ for $t > t_B$. It is continuous, strictly decreasing, and *slow*: $\dot{x}_{\mathcal{E}} > -\lambda N$, on the arc because $A_m < N$ and on the high-quality part by (R2) and Lemma C.1(d). It is also continuously differentiable at the tip t_B: by the implicit function theorem applied to (6.9), whose derivative in y at $y = y_m = 0$ is $\lambda K \neq 0$, $y_H = Y(y_m)$ for a function Y that is smooth near 0 with $Y(0) = 0$ and $Y'(0) = 1$, so $\dot{x}_H = \dot{x}_L + Y'(y_m)\dot{y}_m \to \dot{x}_L(t_B) + \dot{y}_m(t_B) = \dot{x}_m(t_B)$ as $t \downarrow t_B$. Along $\mathcal{E}$ define the boundary gradient $\omega = e_N e^x + rm/\lambda$, which equals $e_N e^{x_m}$ on the arc and Φ_x on the band frontier, and the boundary value $W_{\mathcal{E}}$, equal to W_{arc} on the arc and to $\Phi + \chi_T$ on the band frontier.

Full-effort characteristics are the lines $x + \lambda N t = \xi$. Let $\xi_* = x_m(t_*) + \lambda N t_*$ be the line through the onset point.

Lemma C.2 (Exits)**.** *Assume* (R2)*. Every line $\xi \ge \xi_*$ meets $\mathcal{E}$ exactly once, at a time $t_e(\xi) \ge t_*$ that is continuous and strictly increasing in ξ, with $t_e(\xi_*) = t_*$ and $t_e(\xi) \to \infty$ as $\xi \to \infty$. Before $t_e(\xi)$, the line lies above $\mathcal{E}$.*

Proof. The function $\xi - \lambda N t - x_{\mathcal{E}}(t)$ is continuous and strictly decreasing in t on $[t_*, \infty)$, because $\mathcal{E}$ is slow; it is nonnegative at t_* for $\xi \ge \xi_*$ and tends to $-\infty$. Monotonicity in ξ is immediate. □

For a point (x, t) with $\xi = x + \lambda N t > \xi_*$ and $t < t_e(\xi)$, let $\tau = t_e(\xi) - t$ and let $P_e = (x_e, t_e)$ be the exit point. Define the full-effort value and gain

$$W^F(x,t) = \int_0^\tau e^{-ru}\Big[rm + \lambda N e_N e^x\Big](x - \lambda N u, t + u)\,\mathrm{d}u + e^{-r\tau} W_{\mathcal{E}}(t_e), \tag{C.5}$$

$$G^F(x,t) = \Gamma(\operatorname{expit} x, \rho(t)) = -r\int_0^\tau e^{-ru} S_1(x - \lambda N u, t + u)\,\mathrm{d}u. \tag{C.6}$$

Here Γ is defined in (6.11), and the integral in (C.5) equals $e^x\big[(r(g-s)+\lambda N e_N)\frac{1-e^{-\Lambda\tau}}{\Lambda} - rge^{-\eta t}\frac{1-e^{-L_N\tau}}{L_N}\big] - s(1-e^{-r\tau})$.

Lemma C.3 (Full-effort region)**.** *Assume* (R2). *W^F is continuously differentiable in the open set $\{(x,t) : \xi > \xi_*,\ t < t_e(\xi)\}$, with*

$$W_x^F(x,t) = \int_0^\tau e^{-ru}\big[rm_x + \lambda N e_N e^x\big](x-\lambda N u, t+u)\,\mathrm{d}u + e^{-r\tau}\omega(P_e), \tag{C.7}$$

and its private gain $rm + \lambda(e_N e^x - W_x^F)$ equals G^F. W^F satisfies $rW = W_t + NG - Krm$, the equation (6.4) with $a = 1$ apart from the optimality condition. If $t \geq t_*$, then $G^F(x,t) > 0$.

Proof. Differentiating (C.5) in x, with $\tau_x = 1/(\lambda N + \dot{x}_{\mathcal{E}}(t_e)) > 0$, produces the integral in (C.7) plus $\tau_x e^{-r\tau}[f(P_e) - rW_{\mathcal{E}} + W'_{\mathcal{E}}]$ with $f = rm + \lambda N e_N e^x$. On the arc, (C.4) gives $f - rW_{\mathcal{E}} + W'_{\mathcal{E}} = \lambda e_N e^{x_m}(N - A_m) = (\lambda N + \dot{x}_m)\omega$. On the band frontier, $W'_{\mathcal{E}} = W_t + W_x \dot{x}_H$ with $W = \Phi + \chi_T$, $W_t - rW = Krm$ ((6.4) with $\alpha = 1$, $G = 0$) and $W_x = \omega$, so $f - rW_{\mathcal{E}} + W'_{\mathcal{E}} = Nrm + \lambda N e_N e^x + \omega \dot{x}_H = (\lambda N + \dot{x}_H)\omega$. In both cases the boundary term is $e^{-r\tau}\omega(P_e)$. At a point where the exit type changes (the line through the tip $(t_B, x_m(t_B))$) the two one-sided computations agree, so W_x^F is continuous; W_t^F is obtained in the same way. Substituting (C.7) and using $\omega(P_e) = e_N e^{x_e} + rm(P_e)/\lambda$,

$$rm + \lambda e_N e^x - \lambda W_x^F = \big[rm + \lambda e_N e^x\big](x,t) - e^{-r\tau}\big[rm + \lambda e_N e^x\big](P_e) - \lambda\int_0^\tau e^{-ru}\big[rm_x + \lambda N e_N e^x\big]\,\mathrm{d}u,$$

and writing the difference of the first two terms as the integral of a derivative along the line gives $-r\int_0^\tau e^{-ru} S_1\,\mathrm{d}u$. The equation $rW = W_t + NG - Krm$ follows from (C.5) by differentiation along the line. If $t \geq t_*$, the line stays above $\mathcal{E}$ until t_e; at times in $[t_*, t_B]$ it is above the arc, where $S_1(x_m(t'), t') \leq 0$ by (C.1), and at times after t_B it is above x_H, where $S_1 < 0$ by Lemma C.1(b). Since $\Theta_1 > 0$ after t_* (Lemma B.2), $S_1(\cdot, t')$ is decreasing, so $S_1 < 0$ along the whole segment and $G^F > 0$. □

C.4 The low-quality band

Fix $\xi > \xi_*$ with exit point $P_e = (x_e, t_e)$. Along the line traced backward, let $\Gamma_\xi(\tau) = G^F(x_e + \lambda N\tau, t_e - \tau)$ and $\sigma_\xi(v) = S_1(x_e + \lambda N v, t_e - v)$, for $0 \leq \tau, v \leq t_e$. By (C.6),

$$e^{r\tau}\Gamma_\xi(\tau) = -r\int_0^\tau e^{rv}\sigma_\xi(v)\,\mathrm{d}v. \tag{C.8}$$

Lemma C.4 (Negative gain at t_E)**.** *Assume* (R2). *$G^F(x, t_E) < 0$ whenever it is defined.*

Proof. By (B.2), $\int_0^\infty e^{-\Lambda u}\Theta_1(t_E + u)\,\mathrm{d}u = 0$. Since $\Theta_1(t_E) < 0$ and Θ_1 is increasing, the partial integral $\int_0^\tau e^{-\Lambda u}\Theta_1(t_E + u)\,\mathrm{d}u$ is negative for every $\tau > 0$: if $\Theta_1(t_E + \tau) \leq 0$ its integrand is nonpositive and negative near zero, and otherwise it equals $-\int_\tau^\infty e^{-\Lambda u}\Theta_1(t_E + u)\,\mathrm{d}u < 0$. Both terms of the expanded gain (6.11) are then negative. □

Lemma C.5 (A single zero along each characteristic)**.** *Assume* (R2)*. For every $\xi > \xi_*$, Γ_ξ has a unique zero $\tau_0(\xi)$ in $(0, t_e - t_E)$; $\Gamma_\xi > 0$ on $(0, \tau_0)$ and $\Gamma_\xi < 0$ on $(\tau_0, t_e - t_E]$; and $S_1 > 0$ at the zero.*

Proof. Write $\sigma_\xi(v) = rs - e^{x_e}\vartheta(v)$ with $\vartheta(v) = \Theta_1(t_e - v)e^{\lambda N v}$. Then $\vartheta'(v) = e^{\lambda N v}[\lambda N C_1 - (\lambda N + \eta)C_2 e^{-\eta(t_e - v)}]$, whose bracket decreases in v; so ϑ is increasing and then decreasing, or decreasing. At $v = 0$, $\sigma_\xi(0) = S_1(P_e) < 0$ (by (C.1) on the arc with $t_e > t_*$, and by Lemma C.1(b) on the band frontier), so $e^{x_e}\vartheta(0) > rs$. The set where $e^{x_e}\vartheta > rs$ is therefore an interval $[0, \tau_1)$, and $\sigma_\xi > 0$ on $(\tau_1, t_e]$; $\tau_1 < t_e - t_E$ because $\Theta_1(t_E) < 0$. By (C.8), $e^{r\tau}\Gamma_\xi$ increases on $[0, \tau_1]$ from zero and decreases on $[\tau_1, t_e]$, and it is negative at $t_e - t_E$ by Lemma C.4. Hence the zero exists, is unique, lies in $(\tau_1, t_e - t_E)$, and $\sigma_\xi(\tau_0) > 0$. □

Lemma C.6 (The upper frontier)**.** *Assume* (R1) *and* (R2)*, and let $t_H(\xi) = t_e(\xi) - \tau_0(\xi)$.*

(a) t_H is continuous and strictly decreasing on (ξ_*, ∞), with $t_H(\xi) \to t_*$ as $\xi \downarrow \xi_*$ and $t_H(\xi) \to t_E$ as $\xi \to \infty$. Its inverse defines $x_{Hc}(t) = \xi(t) - \lambda N t$ on (t_E, t_*), with $x_{Hc}(t) \to \infty$ as $t \downarrow t_E$ and $x_{Hc}(t_*) = x_m(t_*)$.

(b) x_{Hc} is continuously differentiable and fast*: $\dot{x}_{Hc} < -\lambda N$.*

(c) For $t \in (t_E, t_*)$, $G^F(x, t) > 0$ if $x > x_{Hc}(t)$, $G^F(x_{Hc}(t), t) = 0$, and $G^F(x, t) < 0$ if $x < x_{Hc}(t)$ and $x + \lambda N t > \xi_*$.

(d) If also (R3) *holds, then $x_{Hc}(t) > x_m(t)$ on (t_E, t_*), and every line $\xi > \xi_*$ lies above the myopic curve between its frontier point and t_*.*

Proof. Continuity. The exit point depends continuously on ξ (Lemma C.2) and $\Gamma_\xi(\tau)$ is jointly continuous; at the zero, $\frac{d}{d\tau}(e^{r\tau}\Gamma_\xi) = -re^{r\tau}\sigma_\xi(\tau_0) < 0$, so the zero moves continuously.

Monotonicity. Let $\gamma(\xi, t) = G^F(\xi - \lambda N t, t)$. At $(\xi, t_H(\xi))$, the derivative along the line is $\gamma_t = rG^F + rS_1 = rS_1 > 0$ (Lemma C.5). The derivative across lines is $\gamma_\xi = G^F_x$. Differentiating (6.11) in log odds x at fixed t, with $\tau_x > 0$,

$$G^F_x = re^x I_1 - r\tau_x e^{-r\tau} S_1(P_e), \qquad I_1 = \int_0^\tau e^{-\Lambda u}\Theta_1(t + u)\,\mathrm{d}u,$$

and at a zero $re^x I_1 = rs(1 - e^{-r\tau_0})$, so $G^F_x = rs(1 - e^{-r\tau_0}) - r\tau_x e^{-r\tau_0} S_1(P_e) > 0$ because $S_1(P_e) < 0$. Hence $dt_H/d\xi = -\gamma_\xi/\gamma_t < 0$, and along the frontier $dx_{Hc}/dt = d\xi/dt_H - \lambda N < -\lambda N$; this is (b).

Limits. For fixed $t > t_E$, $\tau \to \infty$ as $x \to \infty$ and $G^F/e^x \to r\int_0^\infty e^{-\Lambda u}\Theta_1(t+u)\,\mathrm{d}u = r\phi^1(\rho(t)) > 0$ by (B.2); so points with large x lie above the frontier, which forces $t_H(\xi) \to t_E$. At the onset exit $\xi = \xi_*$, $\sigma_{\xi_*}(0) = S_1(x_m(t_*), t_*) = 0$ and, using $e^{x_m}\Theta_1 = rs$ and $e^{x_m}\Theta_1' = s(r + \lambda K + \eta)\lambda A_m(\rho_*)$,

$$\sigma'_{\xi_*}(0) = e^{x_m}(\Theta_1' - \lambda N \Theta_1) = s\lambda\big[(r + \lambda K + \eta)A_m(\rho_*) - Nr\big] > 0$$

by (R1). By continuity, for ξ near ξ_*, $\sigma_\xi(v) \geq \sigma_\xi(0) + k_0 v$ on a fixed interval with $k_0 > 0$ and $\sigma_\xi(0) < 0$, $\sigma_\xi(0) \to 0$; by (C.8), Γ_ξ becomes negative before $\tau = -4\sigma_\xi(0)/k_0 \to 0$. Hence $\tau_0(\xi) \to 0$

and $t_H(\xi) \to t_*$. (If the inequality in (R1) is reversed, the zeros do not converge to the onset point, and the frontier detaches from the arc.)

(c) Let $t \in (t_E, t_*)$ and let $\xi = x + \lambda N t$ be the line through (x,t). Because t_H is decreasing, $t_H(\xi) < t$ if $x > x_{Hc}(t)$, and $t_H(\xi) > t$ if $x < x_{Hc}(t)$ and $\xi > \xi_*$. In the first case $t \in (t_H(\xi), t_e(\xi))$, where $G^F > 0$; in the second $t \in (t_E, t_H(\xi))$, where $G^F < 0$; both by Lemma C.5.

(d) Let $\bar{x}_*(t) = \xi_* - \lambda N t$ be the full-effort line through the onset point. Going backward from t_*, the myopic curve rises at rate $\lambda A_m < \lambda N$ on (t_m, t_*), so $\bar{x}_* > x_m$ there, and $\bar{x}_* > x_m$ on (t_c, t_*) by the definition of ρ_c; in particular $t_c \le t_m$, because $t_E < t_m$ (Lemma B.2). A line $\xi > \xi_*$ lies above $\bar{x}_*$, hence above the myopic curve on $(\max\{t_c, t_E\}, t_*)$. If it crosses the myopic curve at an earlier time $t_\times \in (t_E, t_c)$ (going backward, for the first time), the crossing point is a point of the myopic curve with quality in (E_N, ρ_c) whose full-effort path exits at $P_e(\xi)$; by (R3), $G^F < 0$ there, so $\tau_0(\xi) < t_e - t_\times$ and the frontier point comes before the crossing. In all cases the frontier point and the segment from it to t_* lie above the myopic curve. □

Lemma C.7 (Interior effort in the low-quality band)**.** *Assume* (R1)–(R3). *For $t \in (t_E, t_*)$, write x_{Hc} for $x_{Hc}(t)$ and let $\mathcal{N}(x,t) = rs(x - x_{Hc}) - \Theta_0(t)(e^x - e^{x_{Hc}}) + \lambda K m(x_{Hc}, t)$. Then on $(x_m(t), x_{Hc}]$*

$$\alpha = \frac{\mathcal{N}}{\lambda K m} = 1 - \frac{1}{\lambda K m(x,t)} \int_x^{x_{Hc}} S_1(y,t)\,\mathrm{d}y \tag{C.9}$$

is continuous and strictly increasing, $\alpha(x_{Hc}) = 1$, and α has a unique zero $x_{Lc}(t) \in (x_m(t), x_{Hc})$. In addition, $S_0(x_{Lc}(t), t) > 0$; x_{Lc} is continuous and strictly decreasing, with $x_{Lc}(t) \to \infty$ as $t \downarrow t_E$ and $x_{Lc}(t_*) = x_m(t_*)$.

Proof. Multiplied by $\lambda K m$, the two sides of the second equality in (C.9) are $\mathcal{N}$ and $\lambda K m - \int_x^{x_{Hc}} S_1$; both equal $\lambda K m(x_{Hc}, t)$ at x_{Hc}, and both have derivative $S_0 = \lambda K m_x + S_1$ in x, which proves it. By Lemma C.5, $S_1(x_{Hc}, t) > 0$; if $\Theta_1(t) > 0$ then $S_1(\cdot, t)$ is decreasing, and otherwise $S_1 > 0$ everywhere; either way $S_1(y,t) > 0$ for $y < x_{Hc}$, so $\mathcal{I}(x) = \int_x^{x_{Hc}} S_1$ is positive and decreasing on $x < x_{Hc}$. Since $\rho(t) > E_N > q$, $m(\cdot, t)$ is increasing and positive above $x_m(t)$, so $\alpha_x = (S_1 m + \mathcal{I} m_x)/(\lambda K m^2) > 0$ there, and α is strictly increasing. As $x \downarrow x_m(t)$, $m \to 0^+$ and $\mathcal{I} > 0$, so $\alpha \to -\infty$; hence the unique zero. At the zero, $\mathcal{N} = 0$ gives $\alpha_x = \mathcal{N}_x/(\lambda K m) = S_0/(\lambda K m)$, so $S_0(x_{Lc}, t) > 0$. Differentiating $\mathcal{N}(x_{Lc}(t), t) = 0$,

$$\partial_t \mathcal{N} = -\dot{x}_{Hc}\, S_1(x_{Hc}, t) + \Theta_0'(t)(e^{x_{Hc}} - e^x) + \lambda K e^{x_{Hc}} g\eta(1-\rho) > 0,$$

because $\dot{x}_{Hc} < 0$, $S_1(x_{Hc}) > 0$ and $x < x_{Hc}$; therefore $\dot{x}_{Lc} = -\partial_t \mathcal{N}/S_0 < 0$. As $t \downarrow t_E$, $\Theta_1(t) < 0$ and $x_{Hc} \to \infty$, so $\mathcal{I}(x) \ge rs(x_{Hc} - x) \to \infty$ for fixed x, hence $x < x_{Lc}(t)$ eventually. At t_* both frontiers equal $x_m(t_*)$. □

Value in the band. For $t \in (t_E, t_*)$ and $x_{Lc}(t) \le x \le x_{Hc}(t)$ put $W = \Phi + \chi_C(t)$ with $\chi_C(t) = W^F(x_{Hc}(t), t) - \Phi(x_{Hc}(t), t)$. Since $W_x^F = \Phi_x$ at the frontier (Lemma C.6(c) and Lemma C.3), $\chi_C' =$

$W_t^F - \Phi_t$ there, and (6.4) with $a = 1$ for W^F gives $\chi_C' - r\chi_C = Krm(x_{Hc}) - (r/\lambda)(rsx_{Hc} - \Theta_0 e^{x_{Hc}})$. By (B.4), W then satisfies (6.4) with $a = \alpha$ from (C.9), and $G = 0$ in the band.

Lemma C.8 (Frontiers near certainty)**.** *Assume* (R1)–(R3). *Put*

$$c_E = g\left(1 - \frac{\lambda}{L_N}\right), \qquad d_E = c_E\eta(1 - E_N), \qquad w_E = \log\left(1 + \frac{\lambda K(gE_N - s)}{d_E}\right) > 0.$$

As $\rho \downarrow E_N$, the low-quality frontiers satisfy

$$e^{x_{Hc}(t(\rho))}(\rho - E_N) \longrightarrow \frac{s}{c_E}, \qquad x_{Hc}(t(\rho)) - x_{Lc}(t(\rho)) \longrightarrow w_E, \tag{C.10}$$
$$1 - p_H(\rho) \sim \frac{c_E}{s}(\rho - E_N), \qquad 1 - p_L(\rho) \sim \frac{c_E e^{w_E}}{s}(\rho - E_N).$$

In particular the width of the band in log odds remains bounded, including as it reaches the known-state edge.

Proof. By (B.2), the infinite-horizon integral of Θ_1 along a full-effort characteristic is $\phi^1(\rho) = c_E(\rho - E_N)$. Subtracting its tail in (6.11), with $z = e^x$, $z_e = e^{x_e}$, and $\rho_e = \rho(t_e)$, gives the exact identity

$$\frac{G^F(x,t)}{r} = z\phi^1(\rho) - s + e^{-r\tau}\left[s - z_e\phi^1(\rho_e)\right].$$

On the upper frontier $G^F = 0$ and $x_{Hc} \to \infty$ as $t \downarrow t_E$. Thus the characteristic label and its exit time tend to infinity, $\tau \to \infty$, and z_e tends to the finite full-quality upper-frontier odds. The last term tends to zero, proving the first limit in (C.10).

The edge identity also gives $\Theta_1(t_E) = -c_E\eta(1 - E_N) = -d_E$. Write $w = x_{Hc} - x_{Lc} \geq 0$ and $z_L = e^{x_{Lc}}$. The lower-frontier identity from (C.9), divided by z_L, is

$$\lambda K\left[g\rho - s - \frac{s}{z_L}\right] = \frac{rsw}{z_L} - \Theta_1(t)(e^w - 1).$$

Near t_E, $-\Theta_1(t) \geq d_E/2$, so boundedness of the left side first bounds w. Since $z_L \to \infty$, passing to the limit gives $e^w \to 1 + \lambda K(gE_N - s)/d_E$. Away from t_E, continuity of both frontiers and their common limit at t_* give a global bound on the width. Finally $1 - \operatorname{expit} x \sim e^{-x}$ as $x \to \infty$, which yields the belief-frontier asymptotics. □

C.5 The waiting region

For $x > x_\ell$ let $T(x)$ be the first time at which the waiting path of constant x meets the boundary of the waiting region: $T(x) = x_{Lc}^{-1}(x)$ if $x > x_m(t_*)$, $T(x) = x_m^{-1}(x)$ if $x_m(t_B) \leq x \leq x_m(t_*)$, and $T(x) = x_L^{-1}(x)$ if $x_\ell < x < x_m(t_B)$. Set $T(x) = \infty$ for $x \leq x_\ell$. The waiting region is $\{t < T(x)\}$, and there

$$W(x,t) = e^{-r(T(x)-t)}W\big(x, T(x)\big), \qquad W = 0 \text{ if } x \leq x_\ell. \tag{C.11}$$

Lemma C.9 (Waiting region)**.** *Assume* (R1)–(R3)*. In the waiting region,* $W_x(x,t) = e^{-r(T-t)}\omega(x,T(x))$ *with* $\omega = e_N e^x + rm/\lambda$ *evaluated at the entry point,* $W_t = rW$*, and*

$$G(x,t) = -r\int_t^{T(x)} e^{-r(u-t)} S_0(x,u)\,\mathrm{d}u < 0.$$

Proof. Let $\bar{W}(x) = W(x,T(x))$ be the value at the entry point. Differentiating (C.11), $W_x = e^{-r(T-t)}[\bar{W}'(x) - rT'(x)\bar{W}(x)]$. If the entry point lies on a lower band frontier, $\bar{W}' = W_x + T'W_t$ with $rW - W_t = \alpha(NG - Krm) = 0$ there, because $\alpha = 0$; so $W_x(x,t) = e^{-r(T-t)}W_x(x,T) = e^{-r(T-t)}\omega$. If it lies on the arc, $\bar{W} = W_{\mathrm{arc}} \circ T$ with $T' = 1/\dot{x}_m = -1/(\lambda A_m)$, and (C.4) gives $\bar{W}' - rT'\bar{W} = T'(W'_{\mathrm{arc}} - rW_{\mathrm{arc}}) = e_N e^x = \omega$. In both cases $W_x = e^{-r(T-t)}\omega$, and $W_t = rW$ by (C.11). Then $G = rm(x,t) + \lambda e_N e^x - e^{-r(T-t)}[rm + \lambda e_N e^x](x,T)$, which equals the stated integral because $\frac{d}{du}\{e^{-r(u-t)}[rm + \lambda e_N e^x](x,u)\} = e^{-r(u-t)}r[m_t - rm - \lambda e_N e^x] = e^{-r(u-t)}rS_0$. For $x \le x_\ell$ the same formula holds with $T = \infty$. Finally $S_0(x,u) > S_0(x,T(x)) \ge 0$ for $u < T(x)$, because Θ_0 is strictly increasing and $S_0 \ge 0$ at every entry point: $S_0(x_{Lc}) > 0$ (Lemma C.7), $S_0(x_m(t),t) \ge 0$ for $t \le t_B$ by (C.1), $S_0(x_L) = 0$, and $S_0(x,\infty) \ge 0$ for $x \le x_\ell$. □

C.6 Regularity of the value

Lemma C.10 (Regularity)**.** *Assume* (R1)–(R3)*, and let* W *be defined by* (C.3)–(C.11) *in the respective regions. Then* W *is continuous on* $\mathbb{R}\times[0,\infty)$*, continuously differentiable on* $\mathbb{R}\times(0,\infty)$*, and* $|W(x,t)| \le C(1+e^x)$*. The value* $v = s + (1-p)W$ *extends continuously to* $[0,1]^2$*, with* $v(1,\rho) = u(\rho)$ *and* $v(p,1) = s + (1-p)W_\infty(\operatorname{logit} p)$*, where* $W_\infty = W(\cdot,\infty)$ *is the stationary value at full quality.*

Proof. Inside each region W is continuously differentiable by the explicit formulas. Across each interface, W is continuous by construction, and W_x is continuous because the private gain vanishes on both sides: at the lower frontiers x_{Lc} and x_L by Lemma C.9; at the band frontiers x_{Hc} and x_H because $G^F = 0$ there and $G = 0$ in the bands; and on the arc because $W_x = e_N e^{x_m}$ from below (Lemma C.9) and from above (Lemma C.3, with $\tau \to 0$). On every interface the one-sided limits of $a(NG - Krm)$ agree, because $G = 0$ there and the action is continuous across the interface or, on the arc, $m = 0$; so $W_t = rW - a(NG - Krm)$ is continuous as well. At the onset and tip points every one-sided limit of W_x is $e_N e^{x_m}$ and every one-sided limit of W_t is rW, since $G = m = 0$ there. At $t = t_E$ the waiting region contains a neighborhood of every point with finite x, because $x_{Lc}(t) \to \infty$ as $t \downarrow t_E$.

For the bound, use the formulas directly: $|W^F| \le C(1+e^x)$ by (C.5), since $W_{\mathcal{E}}$ is bounded on $\mathcal{E}$; in the high-quality band x lies in a bounded interval and χ_T is bounded; in the low-quality band $W = W^F(x_{Hc},t) - \int_x^{x_{Hc}} \Phi_x$ and the width $x_{Hc} - x_{Lc}$ is bounded by Lemma C.8; and in the waiting region $|W|$ is at most its value at the entry point, which has the same x.

The edge $p = 1$. Put

$$u^1(t) = \frac{r(g-s) + \lambda N e_N}{\Lambda} - \frac{rg[1-\rho(t)]}{L_N}, \qquad \Phi^1(t) = e_N + \frac{r}{\lambda}[g\rho(t) - s].$$

The first expression is $u(\rho(t)) - s$ on the active branch of (3.3), and the second is the coefficient of e^x in Φ_x; the first claim holds because $rg/\Lambda - rg/L_N = \eta rg/(\Lambda L_N)$. Splitting the integral in (C.5) into its infinite-horizon value and its tail gives

$$W^F(x,t) = e^x u^1(t) - s + e^{-r\tau}\big[W_{\mathcal{E}}(t_e) - e^{x_e}u^1(t_e) + s\big].$$

The bracket is bounded uniformly over exit points, so $W^F = e^x u^1(t) + O(1)$ uniformly throughout the full-effort region. Consequently $(1-p)W^F \to u(\rho) - s$ uniformly as $p \to 1$, including along sequences with $t \to t_E$ or $t \to \infty$.

In the low-quality band, $W(x,t) = W^F(x_{Hc},t) - \int_x^{x_{Hc}} \Phi_x(y,t)\,\mathrm{d}y$ with $\Phi_x(y,t) = e^y\Phi^1(t) - rs/\lambda$. Thus, with $\delta = x_{Hc} - x$,

$$(1-p)W(x,t) = e^{\delta}u^1(t) - (e^{\delta}-1)\Phi^1(t) + O(e^{-x}),$$

uniformly because $0 \le \delta \le x_{Hc} - x_{Lc}$ is bounded by Lemma C.8. A sequence in this band with $x \to \infty$ has $t \to t_E$. Known-state indifference gives $u^1(t_E) = \Phi^1(t_E) = u(E_N) - s$, so the displayed expression converges to $u(E_N) - s$ throughout the band. In the waiting region, $T(x) \to t_E$ as $x \to \infty$, and (C.11) then gives the limiting surplus $e^{-r(t_E - t)}e_N(1-E_N)^{-\beta} = e_N(1-\rho)^{-\beta}$. These estimates prove joint continuity at $p = 1$. At $\rho = 1$ and finite x, x_L, x_H and χ_T converge to their stationary values; the same formulas in each region and their matching boundary values give continuity there. At $p = 0$, $v = s$ throughout a neighborhood of the edge. □

C.7 Proof of Theorem 2

By Lemma B.2, $q < E_N < \rho_* < \rho_B < 1$. Define $a(x,t)$ as in Table 1: zero in the waiting region, α in the bands, one in the full-effort regions, and A_m/N on the arc. For each t, a is nondecreasing in x: it is zero below the lower frontier, increasing in the bands (Lemma C.1(c), Lemma C.7), and one above the upper frontier. Its only discontinuity in x is at $x_m(t)$ for $t \in [t_*, t_B]$, where $m = 0$ and $G = 0$. By Lemma C.10 and the region-by-region verification of (6.4), with the signs of G established in Sections C.1 to C.5, the hypotheses of Theorem 4 hold; at $\rho = 1$ the stationary versions hold by the same lemmas at $t = \infty$. Hence a^* is a symmetric Markov perfect equilibrium with value $v = s + (1-p)W$.

Part (a): the frontiers are $p_L = \operatorname{expit} x_{Lc}$, $p_H = \operatorname{expit} x_{Hc}$ on (E_N, ρ_*), the myopic curve on $[\rho_*, \rho_B]$, and L, $\operatorname{expit} x_H$ on $(\rho_B, 1]$; continuity at ρ_* and ρ_B and monotonicity follow from Lemmas C.1, C.6 and C.7, and $p_H(1) = \bar{p}_N < q$ because $y_H < y_m$. Part (b): monotonicity and the location of the jump were shown above. Off the closed arc a^* is jointly continuous: the frontiers x_{Lc}, x_{Hc}, x_L and x_H are continuous; the band efforts (6.10) and (C.9) are jointly continuous and

tend to 0 and 1 at the lower and upper frontiers of their bands; $x_{Lc}, x_{Hc} \to \infty$ as $t \downarrow t_E$, so $a^* = 0$ near every point (x, t_E); and x_L, x_H and the band effort converge to their stationary values as $t \to \infty$. Part (c) is Lemma C.10 together with the signs of G; for $p \le \ell_N$ players choose zero effort forever, belief remains fixed while quality recovers, and $W = 0$. Part (d): on the arc $gp\rho = s$, so each player's flow is s, and A_m does not involve N. □

D The conditions: proof of Proposition 2

D.1 Condition (R1)

From (4.1), $\rho_* - q = \eta(\rho_B - q)/(\eta + \lambda K)$ and $1 - \rho_* = [\eta(1-\rho_B) + \lambda K(1-q)]/(\eta + \lambda K)$, so

$$A_m(\rho_*) = \frac{\eta(1-\rho_B) + \lambda K(1-q)}{\lambda(\rho_B - q)} > K\frac{1-q}{\rho_B - q} > K.$$

Hence $(r+\lambda K+\eta)A_m(\rho_*) - Nr > (r+\lambda K+\eta)K - (K+1)r = K(\lambda K+\eta) - r$, which is nonnegative if $r \le K(\lambda K + \eta)$. Lemma C.6 shows the role of (R1): $s\lambda[(r + \lambda K + \eta)A_m(\rho_*) - Nr]$ is the rate at which the source S_1 rises along the full-effort path traced backward from the onset point; if this rate is negative, the zeros of the full-effort gain stay away from the onset and the upper frontier of the low-quality band detaches from the sliding arc.

D.2 Condition (R2)

By Lemma C.1(d), (R2) is equivalent to $\dot{x}_H > -\lambda N$ on (t_B, ∞). Write $\epsilon_H = e^{y_H} - 1$. Since $e^y - 1 - y \le e^y - 1$, (6.9) gives $\lambda K e^{y_H - y_m} \ge \lambda K - r\epsilon_H$, so

$$\Psi \ge r\epsilon_H\Big[\lambda N - \frac{\eta(1-\rho)}{\rho - \rho_\Theta}\Big] + \lambda(N - A_m)(\lambda K - r\epsilon_H) = -r\epsilon_H\,\dot{y}_m + \lambda^2 K(N - A_m),$$

where $\dot{y}_m = \dot{x}_m - \dot{x}_L = \eta(1-\rho)(\rho_\Theta - q)/[(\rho - \rho_\Theta)(\rho - q)]$ by (C.2). Since $y_H < y_m$, (C.2) gives $r\epsilon_H < \eta(\rho - \rho_B)/(\rho - q)$. Using $\lambda(N - A_m) = (\lambda N + \eta)(\rho - \rho_m)/(\rho - q)$, $\Psi > 0$ follows from

$$\eta^2(1-\rho)\,\frac{\rho - \rho_B}{\rho - \rho_\Theta}\,\frac{\rho_\Theta - q}{\rho - q} < \lambda K(\lambda N + \eta)(\rho - \rho_m).$$

Both fractions on the left are below one because $q < \rho_\Theta < \rho_B \le \rho$. It therefore suffices that $\eta^2(1-\rho) \le \lambda K(\lambda N + \eta)(\rho - \rho_m)$ for $\rho \ge \rho_B$, and since the left side decreases and the right side increases in ρ, that this holds at ρ_B. Using $(\lambda N + \eta)(1 - \rho_m) = \lambda N(1-q)$, the condition at ρ_B reads

$$\frac{e_N}{\eta h} = 1 - \rho_B \le \frac{\lambda^2 KN(1-q)}{\lambda K(\lambda N + \eta) + \eta^2}.$$

By Lemma B.2, $e_N/(\eta h) < \lambda(1-q)/L_N$, so it suffices that $\lambda K(\lambda N + \eta) + \eta^2 \le \lambda K N L_N$, which simplifies to $\eta^2 \le \lambda K[Nr + K(\lambda N + \eta)]$. If $\eta \le \lambda K$, then $\eta^2 \le \lambda K \eta \le \lambda K \cdot K(\lambda N + \eta)$.

D.3 Condition (R3)

Lemma D.1 (Closed form for (R3))**.** *Let $t \in (t_E, t_c)$, $\rho_t = \rho(t)$, and let the full-effort line through $(x_m(t), t)$ exit $\mathcal{E}$ at t_e, with $\tau = t_e - t$ and $D = x_m(t) - \lambda N\tau - x_m(t_e) \le 0$ ($D = 0$ for an exit on the arc). Then*

$$G^F(x_m(t), t) = -rs\Big[\frac{\lambda J(t)}{\rho_t - q} - e^{-r\tau}\big(1 - e^D\big)\Big], \qquad J(t) = \int_0^\tau e^{-\Lambda_N u}(\rho_{t+u} - \rho_\varepsilon)\,\mathrm{d}u,$$

with $\rho_\varepsilon = q + e_N/g$, and $J(t) = (1-\rho_\varepsilon)\frac{1-e^{-\Lambda_N\tau}}{\Lambda_N} - (1-\rho_t)\frac{1-e^{-L_N\tau}}{L_N}$. Thus (R3) *holds if and only if $\lambda J/(\rho_t - q) > e^{-r\tau}(1 - e^D)$ for every $t \in (t_E, t_c)$.*

Proof. Write $\Theta_1 = \Lambda_N(g\rho - s) - \frac{d}{du}(g\rho) - \lambda(g\rho - s - e_N)$ along $u \mapsto \rho_{t+u}$, using $r + \lambda K = \Lambda_N - \lambda$ and $\frac{d}{du}g\rho = g\eta(1-\rho)$. Then

$$\int_0^\tau e^{-\Lambda_N u}\Theta_1\,\mathrm{d}u = (g\rho_t - s) - e^{-\Lambda_N\tau}(g\rho_{t+\tau} - s) - \lambda g J(t).$$

Multiply by $re^{x_m(t)}$: on the myopic curve $e^{x_m(t)}(g\rho_t - s) = s$, and $e^{x_m(t)}e^{-\lambda N\tau}(g\rho_{t+\tau} - s) = se^D$ because the line is at log-odds distance D from the myopic curve at t_e. Since $e^{x_m(t)} = q/(\rho_t - q)$ and $g/s = 1/q$, (6.11) gives the formula. □

Proof of part (c). Let $F(t) = \lambda N(t_B - t) - \log\frac{\rho_B - q}{\rho_t - q}$. Then $F'(t) = \lambda(A_m(\rho_t) - N) > 0$ on (t_E, t_m), which is nonempty because $E_N < \rho_m$ (Lemma B.2), and (R3a) states $F(t_E) > 0$. Hence $F > 0$ on (t_E, t_m): the full-effort line through every myopic point with $t \in (t_E, t_m)$ is below the myopic curve at t_B. Since $t_c \le t_m$ (proof of Lemma C.6(d)) and the distance between the myopic curve and the full-effort line through the onset point decreases on (t_E, t_m), every myopic point with $t < t_c$ lies on a line $\xi > \xi_*$. That line is above $\mathcal{E}$ at t_* and below it at t_B, so by Lemma C.2 its unique exit lies on the arc, and $D = 0$. By Lemma D.1, (R3) then reduces to $J(t) > 0$ on (t_E, t_c). If $(1 - E_N)^\beta \le \lambda/(\lambda + r)$, then $\rho_\varepsilon \le E_N$: indeed, by (3.1)–(3.2),

$$\frac{\rho_\varepsilon - q}{E_N - q} = \frac{e_N/g}{E_N - q} = \frac{r(1-E_N)^\beta}{\lambda[1 - (1-E_N)^\beta]},$$

and the integrand of J is positive. Otherwise let $k(t) = \int_t^{t_*} e^{-\Lambda_N u}(\rho_u - \rho_\varepsilon)\,\mathrm{d}u$. The second alternative of (R3b) states exactly $k(t_E) \ge 0$, which requires $\rho_\varepsilon < \rho_*$. Since $k(t_*) = 0$ and $k'(t) = -e^{-\Lambda_N t}(\rho_t - \rho_\varepsilon)$ is positive before $t_\varepsilon = -\eta^{-1}\log(1 - \rho_\varepsilon)$ and negative after, $k > 0$ on (t_E, t_*). Finally $e^{-\Lambda_N t}J(t) = k(t) + \int_{t_*}^{t_e} e^{-\Lambda_N u}(\rho_u - \rho_\varepsilon)\,\mathrm{d}u > 0$, because $\rho_u > \rho_* > \rho_\varepsilon$ for $u > t_*$. □

D.4 Large groups (part (d))

Fix the other primitives and let $N \to \infty$. By (3.1), $E_N - q = \eta(1-q)/(\lambda N^2) + o(N^{-2})$, and by (3.2), $e_N = C_E/N^2 + o(N^{-2})$ with $C_E = r\eta h(1-q)^{1+\beta}/\{\lambda[1 - (1-q)^\beta]\}$; hence $1 - \rho_B = O(N^{-2})$ and $\rho_* - q = \eta(1-q)/(\lambda N) + o(N^{-1})$. Eventually $r \le K(\lambda K + \eta)$ and $\eta \le \lambda K$, so (R1) and (R2) hold.

In (R3a), the left side is $2\log N + O(1)$ and the right side is $(\lambda N/\eta)[2\log N + O(1)]$, so (R3a) holds. For (R3b), $1-(1-\rho_\varepsilon)/(1-E_N) = O(N^{-2})$. With $\tau = t_* - t_E = \eta^{-1}\log\frac{1-E_N}{1-\rho_*} = 1/(\lambda N) + o(N^{-1})$, $k_1 = \Lambda_N\tau \to 1$ and $k_2 = L_N\tau = k_1 + \eta\tau$,

$$M(\tau) = \frac{k_1}{k_2}\,\frac{1-e^{-k_2}}{1-e^{-k_1}} = 1-(k_2-k_1)\Big[\frac{1}{k_1} - \frac{1}{e^{k_1}-1}\Big] + O((k_2-k_1)^2) = 1 - \frac{\eta}{\lambda N}\Big(1-\frac{1}{e-1}\Big) + o(N^{-1}),$$

so $1-M(\tau)$ is of order $1/N$ and (R3b) holds for large N. □

D.5 Rates of convergence

Lemma D.2 (Threshold and learning rates)**.** *For each N for which Condition 1 holds (all sufficiently large N, by Proposition 2(d)), consider the equilibrium of Theorem 2, and let $p_{\infty,N}$ be the limiting posterior without a discovery, P_N the probability of a discovery, and I_N total effort on the no-discovery path. Write $P_\infty(p) = (p-q)_+/(1-q)$ and $I_\infty(p) = [\operatorname{logit} p - \operatorname{logit} q]_+/\lambda$. For every initial quality and $0<p<1$,*

$$\begin{aligned}
&0 \le \min\{p,q\} - p_{\infty,N} \le q - \ell_N,\\
&0 \le P_N - P_\infty(p) \le \frac{q-\ell_N}{1-\ell_N},\\
&0 \le I_N - I_\infty(p) \le \frac{1}{\lambda}\log\left(1 + \frac{e_N}{rh(1-q)}\right).
\end{aligned}$$

The three bounds are $O(N^{-2})$. The upper full-quality cutoff satisfies

$$q - \bar p_N \sim \frac{qC_E^2}{2\lambda r h^2(1-q)}\,N^{-5}.$$

Proof. By Theorem 3, $p_{\infty,N} = \min\{p,\ell_N\}$, $P_N = (p-\ell_N)_+/(1-\ell_N)$ and $I_N = [\operatorname{logit} p - \operatorname{logit}\ell_N]_+/\lambda$. The first two bounds follow by considering $p\le\ell_N$, $\ell_N<p<q$ and $p\ge q$. The positive-part map is one-Lipschitz, and

$$q-\ell_N = \frac{qe_N}{rh+e_N}, \qquad y_{m,\infty} = \operatorname{logit} q - \operatorname{logit}\ell_N = \log\left(1+\frac{e_N}{rh(1-q)}\right),$$

which proves the third bound and all three rates.

For the sharper rate of the upper full-quality cutoff, put $y_{H,\infty} = \operatorname{logit}\bar p_N - \operatorname{logit}\ell_N$, the root of (6.9) at full quality, and $\delta_N = y_{m,\infty} - y_{H,\infty}$. Equation (6.9) gives

$$1 - e^{-\delta_N} = \frac{r}{\lambda(N-1)}\left(e^{y_{m,\infty}-\delta_N} - 1 - y_{m,\infty} + \delta_N\right).$$

Since $0<\delta_N<y_{m,\infty} = O(N^{-2})$, the right side is $O(N^{-5})$, hence $\delta_N = O(N^{-5})$ and $\delta_N/y_{m,\infty}\to 0$.

Taylor expansion therefore yields

$$\delta_N \sim \frac{r y_{m,\infty}^2}{2\lambda N} \sim \frac{C_E^2}{2\lambda r h^2 (1-q)^2} N^{-5}.$$

Finally $\bar{p}_N = \operatorname{expit}(\operatorname{logit} q - \delta_N)$ and the derivative of inverse log odds at $\operatorname{logit} q$ is $q(1-q)$. The remaining threshold equivalents in Proposition 4(a) follow from the expansions in the proof of Proposition 2(d). □

D.6 An exact certificate at the baseline

Lemma D.3 (Baseline conditions)**.** *The baseline primitives* $N = 3$, $\lambda = \eta = s = 1$, $h = 2$ *and* $r = 1/5$ *satisfy* (R1)–(R3).

Proof. The sufficient inequalities for (R1) and (R2) are immediate: $r = 1/5 < 6 = K(\lambda K + \eta)$ and $\eta^2 = 1 < 86/5 = \lambda K[Nr + K(\lambda N + \eta)]$. To verify (R3) by part (c) of Proposition 2, put $w = (1 - E_N)^{1/5}$. The known-state root equation becomes

$$470w^6 - 512w^5 - 210w + 231 = 0, \qquad e_N = \frac{10w^6}{21(11 - 10w)}.$$

The polynomial is strictly positive at $8589/10000$ and strictly negative at $859/1000$, as can be checked with integer arithmetic. Uniqueness of the known-state root places w between these rational endpoints. The functions $1 - w^5$ and $10w^6/[21(11 - 10w)]$ are respectively decreasing and increasing there. Substitution of the endpoints gives the convenient rational enclosures

$$\frac{532}{1000} < E_N < \frac{534}{1000}, \qquad \frac{79}{1000} < e_N < \frac{80}{1000}, \qquad \frac{96}{100} < \rho_B < \frac{1921}{2000}.$$

Consequently $(\rho_B - q)/(E_N - q) < 15$, whereas $(1 - E_N)/(1 - \rho_B) > 11$. Since $\lambda N/\eta = 3$, these inequalities imply (R3a).

For (R3b), $\rho_\varepsilon = q + e_N/2 < 27/50$, so

$$\frac{1 - \rho_\varepsilon}{1 - E_N} > \frac{115}{117}.$$

Here $\rho_* = (1 + \rho_B)/3 > 49/75$, and $\tau = t_* - t_E$ satisfies

$$e^\tau = \frac{1 - E_N}{1 - \rho_*} > \frac{233/500}{26/75} > \frac{4}{3} > e^{1/4},$$

where the last inequality follows from $e^x < 1/(1 - x)$ for $0 < x < 1$. Thus $\tau > 1/4$. The function

$$M(T) = \frac{\int_0^T e^{-(16/5)u} e^{-u} \, \mathrm{d}u}{\int_0^T e^{-(16/5)u} \, \mathrm{d}u}$$

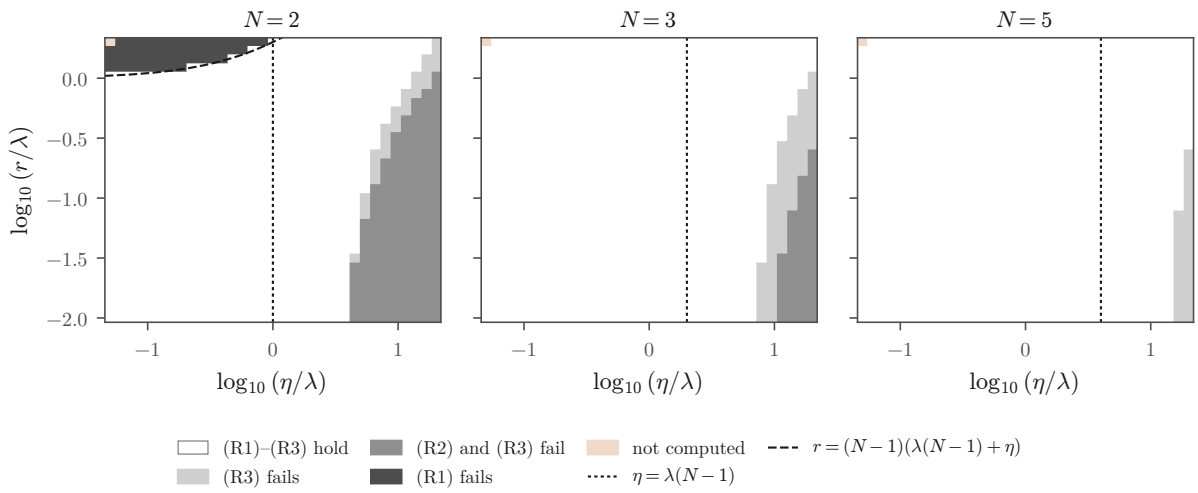


Figure 14: Conditions (R1)–(R3) across primitives ($\lambda = 1$, $q = 0.5$). White cells: all three numerical checks pass. Grey cells: at least one fails, as the legend shows. In the pale cell at the top left of each panel the high-quality band falls below the numerical resolution cutoff. The dotted vertical line is $\eta = \lambda(N-1)$, to the left of which (R2) holds by Proposition 2(b), and the dashed curve is $r = (N-1)(\lambda(N-1)+\eta)$, below which (R1) holds by Proposition 2(a).

decreases in T, since it is the weighted mean of the decreasing function e^{-u} over $[0, T]$. On $[0, 1/4]$, Taylor's bound gives $e^{-u} \le 1 - u + u^2/2 \le 1 - (7/8)u$. Under the probability density proportional to $e^{-(16/5)u}$ on this interval, the mean of u is at least $e^{-4/5}/8 > 1/40$. Hence

$$M(\tau) \le M(1/4) < 1 - \frac{7}{320} = \frac{313}{320} < \frac{115}{117}.$$

This verifies the second alternative of (R3b) and hence (R3) on its entire required interval. □

D.7 A map of the conditions

Figure 14 samples the conditions on a 33×33 grid with $\eta/\lambda \in [0.05, 20]$, $r/\lambda \in [0.01, 2]$ and $q = 0.5$. All three checks pass jointly in 76%, 89% and 97% of cells for $N = 2, 3, 5$, respectively. These are sampled inequalities; analytic guarantees come from Proposition 2. In every resolved cell with $\eta \le \lambda(N-1)$ and $r \le K(\lambda K + \eta)$, all three checks pass. One cell per panel, at $\eta/\lambda = 0.05$ and $r/\lambda = 2$, is unresolved: $\beta = 40$ and $1 - \rho_B \in [2.0, 6.4] \times 10^{-14}$, below the code's band-resolution cutoff of 10^{-12}.

E Proofs for Section 5

E.1 Equilibrium paths and learning (Theorem 3(b) and (d))

Fix an initial state (x_0, t_0) and follow the profile path of a^*, which exists and is unique because a^* is an equilibrium in the sense of Definition 1 (Theorem 2); in this appendix we call it the trajectory.

Barriers. Let $\zeta(t)$ be a lower band frontier (x_{Lc} or x_L). Both are decreasing and effort vanishes on them, so for a trajectory in the band, $\frac{d}{dt}(x-\zeta) \to -\dot{\zeta} > 0$ as $x \downarrow \zeta$: a trajectory in a band never crosses its lower frontier. The upper frontier of the high-quality band satisfies $\dot{x}_H > -\lambda N$ by (R2), while $\dot{x} = -\lambda N\alpha \to -\lambda N$ as $x \uparrow x_H$; hence $\frac{d}{dt}(x_H - x) \to \dot{x}_H + \lambda N > 0$ near the frontier, so trajectories in that band never reach it either, and the band is invariant. The upper frontier of the low-quality band is fast, $\dot{x}_{Hc} < -\lambda N$, so $\frac{d}{dt}(x_{Hc} - x) \le \dot{x}_{Hc} + \lambda N < 0$ along a trajectory in the band: the trajectory leaves the band through x_{Hc} into the full-effort region, unless the band closes first at the onset point, in which case the trajectory reaches the arc.

Order of phases (d). A waiting trajectory keeps x fixed until $T(x_0)$, at which it enters the low-quality band, the arc or the high-quality band, according to the definition of $T(x)$ in Section C.5. From the low-quality band it moves to the full-effort region or to the onset point. From the full-effort region it moves along a line to its exit on $\mathcal{E}$ (Lemma C.2), that is, to the arc or to the high-quality band. On the arc it slides until t_B and then enters the high-quality band, which is invariant. No other transitions occur.

Convergence, part (b). If $x_0 \le x_\ell$ the trajectory waits forever. Otherwise it reaches the high-quality band at a finite time (if $\rho_0 = 1$, it enters the stationary band immediately or after a phase of full effort), after which $x_L(t) < x(t) < x_H(t)$. The trajectory is nonincreasing and bounded below by x_ℓ, so it converges to some $x_\infty \ge x_\ell$. If $x_\infty > x_\ell$, then, since $x_L(t) \to x_\ell$, eventually $y = x - x_L \ge \delta > 0$ and (6.10) gives effort at least $r(e^{\delta} - 1 - \delta)/(\lambda K) > 0$, so $x \to -\infty$, a contradiction. Hence $x_\infty = x_\ell$ and $p_\infty = \ell_N$. Total effort is $\int A^* = (x_0 - x_\infty)/\lambda$, and the probability of a discovery is $p_0(1 - Q_\infty) = p_0(1 - e^{x_\infty - x_0}) = (p_0 - \ell_N)/(1 - \ell_N)$. Since $x(t) > x_L(t) \ge x_\ell$ for every t (at $\rho = 1$, $\dot{y} \approx -\lambda N r y^2/[2\lambda K(1 - e^{-y_m})]$ near $y = 0$, so y stays positive), $\int_T^\infty A^* = (x(T) - x_\ell)/\lambda > 0$ for every T. □

E.2 The tail of discovery (Remark 1)

By the argument in Section E.1, an unsuccessful path with $\ell_N < p_0 < 1$ enters the high-quality band after a finite time; here t is calendar time. Write $x_\ell = \operatorname{logit} \ell_N$, $\hat{y}(t) = \operatorname{logit} p_t - x_\ell > 0$ and $y_L(t) = \operatorname{logit} L(\rho_t) - x_\ell$. If initial quality is below one, (6.7) and $1 - \rho_t = (1 - \rho_0)e^{-\eta t}$ give $y_L(t) = O(e^{-\eta t})$ and $y_m(\rho_t) = y_{m,\infty} + O(e^{-\eta t})$; at full quality both differences vanish identically. The distance inside the band is $y(t) = \hat{y}(t) - y_L(t) \in (0, \hat{y}(t)]$. Once $\hat{y}$ is small, the denominator in (6.10) is bounded away from zero, and $e^y - 1 - y \le Cy^2$ for a constant C. Therefore

$$0 \le -\dot{\hat{y}}(t) = \lambda N \alpha(p_t, \rho_t) \le C' \hat{y}(t)^2.$$

Integrating gives $\hat{y}(t) \ge 1/(C't + C'')$ for large t, so $y_L(t)/\hat{y}(t) \to 0$. Substituting in (6.10) now yields

$$\frac{-\dot{\hat{y}}(t)}{\hat{y}(t)^2} \longrightarrow \frac{Nr}{2K(1 - e^{-y_{m,\infty}})} = c_N.$$

Since $(1/\hat{y})' = -\dot{\hat{y}}/\hat{y}^2$, integration gives $\hat{y}(t) \sim 1/(c_N t)$, and hence $p_t - \ell_N \sim \ell_N(1-\ell_N)/(c_N t)$. By Bayes' rule, the probability that a discovery occurs after t is exactly

$$\mathbb{P}(t < \tau < \infty) = \frac{(1-p_0)(p_t - \ell_N)}{(1-p_t)(1-\ell_N)},$$

the probability $(1-p_0)/(1-p_t)$ of no discovery by t times the probability $(p_t - \ell_N)/(1-\ell_N)$ of a later discovery; this proves (5.2). Dividing by the positive probability of an eventual discovery and integrating the conditional survival function, whose tail is of order $1/t$, gives $\mathbb{E}[\tau \mid \tau < \infty] = \infty$. □

E.3 Distortions (Theorem 3(a) and (c))

Part (a). For $\rho \in (E_N, R_N)$ and $p > p_H(\rho)$, Theorem 2 gives full effort (also on $[\rho_*, \rho_B]$, where $p_H = q/\rho$ and the arc is excluded), and at $p = 1$ the known-state equilibrium uses full effort because $\rho > E_N$. By Theorem 1, the planner waits at every belief when $\rho \le R_N$.

Part (c), first claim. For $p < \ell_N$, $\operatorname{logit} p < x_\ell$, so the equilibrium never experiments from (p, ρ). The planner's frontier p^P is continuous and decreasing with $p^P(1) = p_N^P < p$, so $p > p^P(\rho)$ for all ρ above some $\bar{\rho}(p) < 1$, and there the planner uses full effort.

Part (c), second claim. For $p_N^P < p_0 < 1$, both no-discovery effort integrals are finite; subtract (5.1) from (3.6). At $p_0 = 1$ both integrals are infinite, which is why the statement concerns uncertain priors. □

E.4 Welfare (Proposition 3)

Let A_t^* be equilibrium aggregate effort along the no-discovery trajectory, a measurable function of time (on the arc, $A_t^* = A_m(\rho_t)$). Consider the planner's policy that uses A_t^* until the first opportunity and the efficient known-state policy afterward. The planner's objective is the average payoff, which depends on effort only through the aggregate; before the first opportunity this policy and the equilibrium induce the same distribution of the first opportunity and the same average flow payoffs, and after it the planner obtains $s + D_N$ instead of $s + e_N$. Hence $V \ge v^* + \mathbb{E}[e^{-r\tau}\mathbf{1}_{\{\tau<\infty\}}](D_N - e_N)$, with $D_N > e_N$ by Section A.1. If $p > \ell_N$ or $p = 1$, a discovery occurs with positive probability, so $V > v^*$. If $p_N^P < p \le \ell_N$, then $v^* = s < V$ by Theorem 2(c) and Section A.5. If $p \le p_N^P$, then $V = s = v^*$. □

E.5 Many players (Proposition 4)

(a) The asymptotics of e_N and E_N were derived in the proof of Proposition 2(d); the sharper rates for learning and for the upper full-quality cutoff are established in Lemma D.2. These give $\rho_B = 1 - e_N/(\eta h) \to 1$, $\rho_* \to q$ and $\ell_N \to rs/(\lambda r h) = q$. Since $\ell_N < \bar{p}_N < q$, also $\bar{p}_N \to q$. (b) is Proposition 2(d). (c) A compact set $\mathcal{C}$ as stated has qualities in some $[\rho_1, \rho_2] \subset (q, 1)$; for N large, $\rho_* < \rho_1$ and $\rho_2 < \rho_B$, and on $[\rho_*, \rho_B]$ the equilibrium is the switch at the myopic curve with sliding

effort A_m. (d) follows from Theorem 3, Theorem 1 and Lemma A.2. □

F Equilibria in deterministic schedules

This appendix shows that equilibria in deterministic schedules exist for all primitives. From a given initial state, players commit to deterministic effort schedules until the first opportunity and use the known-state equilibrium afterward. After a deviation in effort, opponents continue their schedules. Such profiles are Nash equilibria from the initial state. The schedules are functions of calendar time rather than of the state, and they need not be optimal after off-path histories, so they are not Markov perfect equilibria in the sense of Definition 1.

F.1 Survival representation

For a schedule a_i let $y_i(t) = \exp(-\lambda \int_0^t a_i)$, $Q = \prod_j y_j$, $Q_{-i} = \prod_{j \neq i} y_j$ and $\rho(t) = 1 - (1 - \rho_0)e^{-\eta t}$. Every schedule corresponds to a path in

$$\mathcal{Y} = \{y : \ y(0) = 1, \ y \text{ locally absolutely continuous}, \ -\lambda y \leq y' \leq 0 \text{ a.e.}\}.$$

For $0 < p_0 < 1$, put $b(t) = g\rho(t) - s$, $c = rs(1 - p_0)/p_0$ and $\gamma_{-i} = Q'_{-i}b + Q_{-i}(g\rho' - rb - \lambda e_N)$. Splitting payoffs at the first opportunity and integrating by parts, as in Lemma A.1, gives

$$J^i - s = p_0\Big[e_N + \frac{r}{\lambda}b(0)\Big] + \frac{rp_0}{\lambda}\int_0^\infty e^{-rt}\{\gamma_{-i}(t)y_i(t) + c\log y_i(t)\}\,\mathrm{d}t, \tag{F.1}$$

which is strictly concave in y_i.

Theorem 5 (Nash equilibrium in schedules)**.** *For all primitives and every initial state, there is a pure symmetric Nash equilibrium in which players follow a common deterministic schedule until the first opportunity and the known-state equilibrium afterward. It is optimal against every adapted unilateral deviation from the initial state.*

Proof. Fix $\delta > 0$ and the Banach space X_δ of continuous f with $e^{-\delta t}f(t) \to 0$ and norm $\sup_t e^{-\delta t}|f(t)|$. The set $\mathcal{Y}$ is nonempty, convex and compact in X_δ: its elements satisfy $e^{-\lambda t} \leq y \leq 1$ and $|y'| \leq \lambda$, so a diagonal argument and the common Lipschitz bound give uniform convergence on compacts, the weighted norm controls the tails, and the differential inequalities pass to limits. The payoff (F.1) is jointly continuous on $\mathcal{Y}^N$: the only delicate term, $\int e^{-rt}Q'_{-i}by_i$, converges because $Q_{-i,n} \to Q_{-i}$ locally uniformly with derivatives bounded by λK, which implies weak-star convergence of those derivatives against every L^1 test function. The varying factor $e^{-rt}by_{i,n}$ converges in L^1 by local uniform convergence and the common discounted tail bound, so this term converges as well. The logarithmic term is controlled by $|\log y| \leq \lambda t$. Strict concavity gives a unique best response, which is continuous in the opponents' paths by compactness and uniqueness. The map $y \mapsto$ (best response to $N - 1$ copies of y) is a continuous self-map of a nonempty compact

convex subset of a Banach space, so it has a fixed point by Schauder's theorem (Schauder, 1930). Conditional on independent randomization, a unilateral adapted deviation specifies one schedule before discovery, which the fixed point beats, and after discovery the known-state equilibrium is optimal. At $p_0 \in \{0, 1\}$ the result is immediate. □

Theorem 6 (Learning in schedule equilibria)**.** *In every Nash equilibrium in deterministic schedules with the known-state continuation, including asymmetric ones, and for every uncertain prior, $p_\infty = \min\{p_0, \ell_N\}$, total effort is $[\operatorname{logit} p_0 - \operatorname{logit} \ell_N]_+/\lambda$, and the probability of a discovery is $(p_0 - \ell_N)_+/(1 - \ell_N)$. If $p_0 > \ell_N$ and $\rho_0 < 1$, effort is positive at arbitrarily late dates.*

Proof. Every no-discovery tail is a Nash equilibrium against the remaining schedules: a profitable tail deviation could be spliced in on the positive-probability event of no discovery. For $T \geq 0$, let J_T^i be player i's continuation payoff at date T conditional on no opportunity before T, and let $S_T(t) = (1 - p_T)/(1 - p_t)$ be the conditional probability of no opportunity in $[T, t)$. A player can always secure at least s by using zero effort and taking the prescribed continuation if opponents discover, so $J_T^i \geq s$. Summing the tail payoffs therefore gives

$$\sum_i (J_T^i - s) = \int_T^\infty e^{-r(t-T)} S_T(t) A_t \{\lambda p_t (rh\rho_t + N e_N) - rs\} \, \mathrm{d}t \geq 0.$$

Put $p_A = rs/[\lambda(rh + N e_N)]$. If $p_T \leq p_A$, the integrand is nonpositive thereafter, and positive total effort would strictly lower the posterior and make the integral negative. Thus effort stops once that bound is reached. Bayes' rule now implies finite total effort and $Q_\infty > 0$; otherwise the posterior would fall below $p_A > 0$.

Each player's tail surplus is at most $(rg + \lambda e_N) \int_T^\infty A$, which tends to zero. If $p_\infty > \ell_N$, at sufficiently late dates a short burst of full effort has strictly positive net coefficient $\lambda p(rh\rho + e_N) - rs$ throughout the burst. Its duration can be chosen fixed and small enough to preserve this strict inequality, since the remaining opponents' total effort tends to zero and $\rho \to 1$. Using zero effort afterward yields a surplus bounded below by a positive constant, a contradiction. Hence $p_\infty \leq \ell_N$.

For the reverse inequality define the marginal tail return

$$M_i(t) = \int_t^\infty e^{-r(u-t)} [\gamma_{-i}(u) y_i(u) + c] \, \mathrm{d}u.$$

A bounded variation δa_i changes the payoff to first order by $-rp_0 \int_0^\infty e^{-rt} M_i(t) \delta a_i(t) \, \mathrm{d}t$. Consequently $M_i > 0$ forces $a_i = 0$, and $M_i < 0$ forces $a_i = 1$, almost everywhere. The function M_i is continuous. Since all survival paths converge and their remaining total variation tends to zero, dominated convergence in the nonderivative terms, together with $\int_t^\infty |Q'_{-i}| \to 0$, gives

$$M_i(t) \longrightarrow Q_\infty \left[\frac{s}{p_\infty} - g - \frac{\lambda e_N}{r} \right].$$

If $p_\infty < \ell_N$, this limit is positive for every player; all effort is therefore zero after some finite date.

If total effort were positive, let $T > 0$ be the essential supremum of its support. After T every y_i is constant. Integrating the quality-recovery term explicitly then gives, for every i,

$$M_i(T) = Q_\infty \left[\frac{s}{p_\infty} - g\rho(T) - \frac{\lambda e_N}{r} \right] > 0.$$

Continuity and the finiteness of the player set imply zero effort on a common interval just before T, contradicting its definition. Thus positive total effort entails $p_\infty = \ell_N$; if there is no effort, $p_\infty = p_0$. Combined with the upper bound, these observations give $p_\infty = \min\{p_0, \ell_N\}$.

If $p_0 > \ell_N$ and $\rho_0 < 1$, a finite last date of effort would instead give $M_i(T) = Q_\infty g[1 - \rho(T)] > 0$, with the same contradiction. Effort consequently has support at arbitrarily late dates. The formulas for total effort and discovery probability follow from Bayes' rule. □

At full quality the schedule equilibrium is explicit: for $\ell_N < p_0 < 1$ all players use full effort until $T_N = [\operatorname{logit} p_0 - \operatorname{logit} \ell_N]/(\lambda N)$ and stop; this is the unique schedule equilibrium, even among asymmetric ones (the marginal-return integrand $c - [r(g-s) + \lambda e_N]Q - \lambda A_{-i}(g-s)Q$ is negative before the posterior reaches ℓ_N). It reaches ℓ_N in finite time, whereas the Markov perfect equilibrium approaches ℓ_N only asymptotically through the band of Theorem 2. Both have the same total effort, but the schedule equilibrium gives every player a strictly higher payoff. Indeed, in terms of cumulative aggregate effort X, with $t(X)$ the date at which it is reached, each player's surplus in a symmetric profile at full quality is

$$\int_0^{X_\infty} e^{-rt(X)} \Big[p_0 e^{-\lambda X} \Big(\frac{r(g-s)}{N} + \lambda e_N \Big) - \frac{rs(1-p_0)}{N} \Big] \, \mathrm{d}X,$$

whose bracket is positive up to X_∞ (where it equals $p_0 Q_\infty \lambda e_N (N-1)/N$); capacity implies $t(X) \geq X/N$, with equality only under full effort. At the baseline with $p_0 = 0.7$ the two payoffs are 1.04590 and 1.04406. Markov perfection thus slows experimentation at full quality: in the band of interior effort every player is indifferent, and aggregate effort is only as high as indifference allows, as in the symmetric equilibrium of Keller et al. (2005).

G Computation and numerical verification

The Python package `expmpe` and script `run_all.py` rebuild every figure and table from the formulas and differential equations of Sections A and C, using NumPy, SciPy and Matplotlib. The README describes reproduction, modules and cached outputs. Two standard-library scripts check the exact rational inequalities of Lemma D.3 and, in 80-digit arithmetic, the asymptotic constants of Lemma D.2. Only the check in Section G.2 uses a discretized game. Table 2 reports the baseline quantities.

Table 2: Baseline primitives and derived quantities

Primitives	$N,\ \lambda,\ h,\ s,\ \eta,\ r$	3, 1, 2, 1, 1, 0.2
	$q = s/(\lambda h)$, $\beta = r/\eta$	0.500, 0.200
Known state	equilibrium's known-state threshold E_N, surplus e_N	0.5326, 0.0793
	planner's known-state threshold R_N, surplus D_N	0.7238, 0.1017
Equilibrium	onset and end of sliding ρ_*, ρ_B	0.6535, 0.9604
	ρ_m (where $A_m = N$), ρ_c	0.6250, 0.5992
	sliding effort at onset and end, $A_m(\rho_*)$, $A_m(\rho_B)$	2.2584, 0.0861
	individual abandonment belief ℓ_N, upper full-quality cutoff $\bar{p}_N$	0.4173, 0.4985
Planner	planner's abandonment belief p_N^P	0.2837
	junction (ρ_c^P, p_c^P)	(0.8422, 0.6130)
Conditions	(R1): $(r + \lambda K + \eta)A_m(\rho_*) - Nr$	6.627
	(R2): sampled $\min_\rho \Psi(\rho)/(\lambda^2 KN)$	0.971
	(R3): sampled $\min_\rho [\lambda J/(\rho - q) - e^{-r\tau}(1 - e^D)]$	0.100

G.1 Computing the equilibrium

Edges and full effort. Brent's method gives E_N and R_N from (3.1) and (3.4). Full-effort exit times solve $\xi - \lambda N t - x_{\mathcal{E}}(t) = 0$; other edge quantities and full-effort values follow their closed forms.

High-quality band and arc. On a quality-clock grid clustered geometrically near t_B, Brent's method gives y_H from (6.9), and adaptive quadrature gives χ_T from (C.3) after subtracting the limit of ϖ. Both use monotone cubic interpolation in $\log(t - t_B)$. Equation (C.4) is integrated backward from t_B to t_* by eighth-order Runge–Kutta with relative tolerance 10^{-12}.

Low-quality band. For exits parametrized along the arc and the upper frontier of the high-quality band, Brent's method finds the first zero τ_1 of σ_ξ and then the zero $\tau_0 > \tau_1$ of Γ_ξ (Lemma C.5), giving (t_H, x_{Hc}) and, by value matching, χ_C. Interpolation in $\log(t - t_E)$ resolves $x_{Hc} \to \infty$ as $t \downarrow t_E$; the lower frontier solves $\mathcal{N} = 0$. Strict monotonicity of t_H is checked on the grid, which stops at $t_H - t_E \simeq 4.95 \times 10^{-10}$ in the baseline. Requests above its resolved belief range are flagged as unresolved; all reported diagnostics lie within that range.

Waiting and paths. Waiting values use (C.11); no-discovery paths use backward Euler on the completed graph of a^*:

$$x_{k+1} \in x_k - \lambda N \Delta \left[a^*(x_{k+1}-, t_{k+1}), a^*(x_{k+1}+, t_{k+1})\right].$$

Monotonicity gives a unique solution, found by bisection: ordinary implicit Euler at continuity points, and effective sliding effort on the arc.

Planner. Closed-form frontiers and survival paths use Brent's method for T_b and the largest root of (A.6); the value (A.2) uses adaptive quadrature.

G.2 An independent check: best responses on a grid

We compute a player's best response by dynamic programming against the constructed opponents' feedback on a finite grid. Time and log odds have the same step Δ. Opponents' feedback is held at its

current-state value within each step, so Δ also plays the role of the update interval of Proposition 6, and the continuation value is linearly interpolated. At each node the deviator maximizes over the full interval of own efforts $[0, 1]$: the one-step objective is piecewise affine, so its maximum is attained at an endpoint or at a control that sends the next state to an interpolation node. The implementation checks every such control. The terminal quality clock is $\max\{t_B + 12/\eta, 14/\eta\}$; its continuation value is the constructed value. Policy evaluation uses the same state grid and terminal condition.

Table 3 reports maxima over sampled beliefs approximately in $[0.076, 0.99988]$ and eleven quality slices between 0.3 and 0.99. The grid best-response gain and the difference between the evaluated and constructed values decrease approximately linearly over the five reported mesh sizes. Proposition 6 concerns the sampled game with a continuous state and gives no rate, and it does not cover the errors from the state interpolation or from the terminal approximation used here.

Table 3: Numerical verification at the baseline

Grid best responses (payoff units)		
Grid step Δ	largest grid best-response gain	largest error of evaluated value
0.016	1.5×10^{-4}	3.5×10^{-3}
0.008	6.9×10^{-5}	1.7×10^{-3}
0.004	3.6×10^{-5}	8.7×10^{-4}
0.002	1.8×10^{-5}	4.3×10^{-4}
0.001	9.2×10^{-6}	2.2×10^{-4}
Regularity and the HJB equation		
largest two-sided difference of W (13 points)		3.2×10^{-9}
largest one-sided derivative mismatch		4.4×10^{-8}
largest HJB residual, waiting		1.8×10^{-11}
largest HJB residual, low-quality band		8.6×10^{-7}
largest HJB residual, full effort		4.0×10^{-10}
largest HJB residual, high-quality band		3.1×10^{-8}
closed-form gain minus finite-difference gain		3.3×10^{-7}

The same table reports two further checks: the two-sided differences of W and the mismatch between one-sided estimates of W_x at thirteen interface points, and the residual of the HJB equation (6.4) evaluated by central differences at a grid of points in every region. The closed-form private gain agrees with its finite-difference counterpart to the reported precision. At full quality the evaluator uses the closed-form values of Section C.1, so the check there only confirms the implementation of those formulas.

H Hidden failures: proofs for Section 8

Throughout this appendix failures are hidden. Besides the log odds $x = \operatorname{logit} p$, we use the *odds* $X = p/(1-p)$ and the *quality-weighted odds* $Y = p\psi/(1-p)$, which satisfy $0 \le Y \le X$. In these

units the myopic surplus of Section 6.1, with expected quality in place of quality, is affine,

$$m = \frac{gp\psi - s}{1-p} = gY - s(1+X), \tag{H.1}$$

and the myopic curve is the line $gY = s(1+X)$. Above the curve $p\psi > q$, so $\psi > q$. The condition $(N-1)s \geq \eta h$ is equivalent to $\lambda(N-1)q \geq \eta$, and $Ns > \eta h$ to $\lambda N q > \eta$.

H.1 The filter (Lemma 8.1)

Proof of Lemma 8.1. Measure time from the date at which the state is (p, ψ). Conditional on $\theta = 1$, quality follows a piecewise-deterministic Markov process. It recovers according to (2.1), opportunities arrive at the rate λA_t whatever quality is, and an opportunity at quality ρ is a success with probability ρ, which stops the process, and otherwise resets quality to zero. Let E_t be the event that $\theta = 1$ and no success has occurred in $(0, t]$, let Z_t be its probability, and let M_t be the expectation of ρ_t times the indicator of E_t, so that $Z_0 = p$ and $M_0 = p\psi$. On E_t, over a short interval $(t, t+\delta]$, a success occurs with probability $\lambda A_t M_t \delta + o(\delta)$, a failure occurs with probability $\lambda A_t (Z_t - M_t)\delta + o(\delta)$ and sends quality to zero, and otherwise quality rises by $\eta(1-\rho_t)\delta + o(\delta)$. Dynkin's formula (Davis, 1993, Ch. 2), applied to the process (t, ρ_t) stopped at the first success, whose jump rate λA_t is bounded and measurable, and to the functions equal to 1 and to ρ before the first success and to zero after it, makes this precise and gives, for almost every t,

$$\dot{Z} = -\lambda A M, \qquad \dot{M} = \eta(Z - M) - \lambda A M. \tag{H.2}$$

An unproductive resource has no success, so the probability of no success in $(0, t]$ is $S_t = Z_t + 1 - p$, and Bayes' rule gives $p_t = Z_t / S_t$ and $\psi_t = M_t / Z_t$, where $Z_t > 0$ because $p > 0$. Hence $1 - p_t = (1-p)/S_t$, so that $X_t = Z_t/(1-p)$ and $Y_t = M_t/(1-p)$, and (H.2) becomes

$$\dot{X} = -\lambda A Y, \qquad \dot{Y} = \eta(X - Y) - \lambda A Y. \tag{H.3}$$

Since $p = X/(1+X)$ and $\psi = Y/X$, differentiation gives (8.1). The probability of a success in $(t, t + \mathrm{d}t]$, given none before t, is $\lambda A_t M_t \,\mathrm{d}t / S_t = \lambda A_t p_t \psi_t \,\mathrm{d}t$. A success shows everybody that $\theta = 1$ and that quality is zero, so it sends the state to $(1, 0)$. □

By (H.3), $X - Y = (1-\psi)p/(1-p)$ falls at the rate η whatever the effort. When failures are observed, $Y = \rho X$ and the belief follows (2.2), so that $\dot{X} = -\lambda A X$, and the second equation of (H.3) still holds. Before a discovery, the two models therefore differ only in which opportunities end the search for a productive resource. Every opportunity does so when failures are observed, and only the successes do so when failures are hidden. After a discovery they also differ in how effort moves quality, and that difference drives Proposition 7.

H.2 Trajectories and the reduction

Let $0 < p < 1$ and $\psi < 1$, and let $c = \log(X - Y) = x + \log(1 - \psi)$. By (H.3), $\dot{c} = -\eta$ whatever the effort, and $\dot{x} = \dot{X}/X = -\lambda A \psi$, with $\psi = 1 - e^{c-x}$. The clock c thus plays the role of the quality clock of Section 6.1, and effort moves only x, at the rate $\lambda\psi$ per unit. If $\psi_0 < 1$, then $c_t = c_0 - \eta t$ is finite and $\psi_t < 1$ for all t. If $\psi_0 = 1$, expected quality stays at one until a success, whatever the effort, and the belief follows (2.2).

We define trajectories as in Section 7.1, with this clock in place of the quality clock. Fix the opponents' strategies and a measurable schedule u of player i, put $c(t) = c_0 - \eta t$ and $\psi(y, c) = \max\{0, 1 - e^{c-y}\}$, and let $\mathcal{A}_{-i}(x, t)$ be the set (7.2) computed from the opponents' aggregate effort $\sum_{j \neq i} a_j(\operatorname{expit} y, \psi(y, c(t)))$ at the clock value $c(t)$. A trajectory is an absolutely continuous function $x(\cdot)$ such that, for almost every t,

$$\dot{x}(t) = -\lambda\, \psi(x(t), c(t)) \big[u(t) + A_{-i}(t)\big] \quad \text{for some } A_{-i}(t) \in \mathcal{A}_{-i}(x(t), t). \tag{H.4}$$

Successes of player i and of the opponents arrive with intensities $\lambda u\, p\psi$ and $\lambda A_{-i}\, p\psi$. At $\psi = 1$, trajectories are those of Section 7.1 at $\rho = 1$. On the edge $p = 1$ the state is ψ, and trajectories are defined in the same way for $\dot{\psi} = (1 - \psi)\big(\eta - \lambda\psi[u + A_{-i}]\big)$, with the opponents' aggregate effort a function of ψ. Profile paths and well-posedness are defined as in Section 7.1.

Lemma H.1 (Well-posedness)**.** *(a) If, at each value of the clock, every opponent's strategy is nondecreasing in x, then from every state with $0 < p < 1$ and $\psi < 1$ the profile is well posed against unilateral deviations before a discovery, and two trajectories generated by schedules u and $\tilde{u}$ from the same state satisfy $|x(t) - \tilde{x}(t)| \leq \lambda \int_0^t |u - \tilde{u}|\, \mathrm{d}\sigma$.*

(b) If every opponent's strategy on the edge $p = 1$ is nondecreasing in ψ, then the profile is well posed against unilateral deviations on the edge.

Proof. (a) The function $\psi(y, c)$ is continuous and nondecreasing in y, so the right side of (H.4) is the Filippov regularization of a Borel function bounded by λN, and trajectories exist as in the proof of Lemma 7.1. They stay in the state space, because $\psi = 0$ where $x \leq c$, so that $\dot{x} = 0$ there while $\dot{c} = -\eta$. Let x and $\tilde{x}$ be trajectories for u and $\tilde{u}$ from the same state, with effective efforts A_{-i} and $\tilde{A}_{-i}$. Both face the same clock path $c(t)$, because effort does not move the clock. If $x(t) > \tilde{x}(t)$, then $A_{-i}(t) \geq \tilde{A}_{-i}(t) \geq 0$, as in the proof of Lemma 7.1, and $\psi(x(t), c(t)) \geq \psi(\tilde{x}(t), c(t)) \geq 0$, so

$$\psi(x)\,(u + A_{-i}) - \psi(\tilde{x})\,(\tilde{u} + \tilde{A}_{-i}) \geq \psi(x)\,(u - \tilde{u}) \geq -|u - \tilde{u}|.$$

Hence $\frac{d}{dt}(x - \tilde{x})^2 \leq 2\lambda |x - \tilde{x}|\, |u - \tilde{u}|$, and the bound and uniqueness follow as in the proof of Lemma 7.1.

(b) Existence is as in (a). Put $\varphi(\psi) = \psi(1 - \psi)$, which has slope at most one in absolute value on $[0, 1]$, and let $\psi > \tilde{\psi}$ be the states at date t of two trajectories with own efforts b and $\tilde{b}$. Then

$A_{-i} \geq \tilde{A}_{-i}$, so

$$\varphi(\psi)(b + A_{-i}) - \varphi(\tilde{\psi})(\tilde{b} + \tilde{A}_{-i}) \geq -\varphi(\psi)|b - \tilde{b}| - N(\psi - \tilde{\psi}),$$

and $\frac{d}{dt}|\psi - \tilde{\psi}| \leq \lambda N|\psi - \tilde{\psi}| + \frac{\lambda}{4}|b - \tilde{b}|$ almost everywhere. By Grönwall's inequality, $|\psi - \tilde{\psi}|(t) \leq \frac{\lambda}{4}\int_0^t e^{\lambda N(t-\sigma)}|b - \tilde{b}|\,\mathrm{d}\sigma$, which gives uniqueness when $b = \tilde{b}$. □

Lemma H.2 (Reduction)**.** *Fix a well-posed profile, an initial state with $0 < p < 1$, and a deviation by player i with no-success schedule u, along whose trajectory the opponents' effective aggregate effort is A_{-i}. If player i's continuation value after a discovery is at most $s + \bar{e}$, then $J^i - s \leq (1-p)W^i$, where, with m_t the myopic surplus (H.1) at date t,*

$$W^i = \int_0^\infty e^{-rt}\Big[\, r\, u_t\, m_t + \lambda \bar{e}\, \psi_t e^{x_t}(u_t + A_{-i,t})\Big]\,\mathrm{d}t,$$

with equality if the continuation value is $s + \bar{e}$.

Proof. The proof of Lemma 6.1 in Section B applies, with the probability of no success $S_t = (1-p)/(1-p_t)$ from the proof of Lemma 8.1, the success intensity $\lambda A_t p_t \psi_t$, and the expected prize rate $g\, u_t p_t \psi_t$ in place of $g\, u_t p_t \rho_t$. □

H.3 After a discovery

On the edge $p = 1$ the state is ψ. Between successes $\dot{\psi} = (1-\psi)(\eta - \lambda A\psi)$, a success arrives at the rate $\lambda A\psi$ and sends ψ to zero, and player i's payoff is $\mathbb{E}\int_0^\infty re^{-rt}[s + a_t^i m_1(\psi_t)]\,\mathrm{d}t$, with $m_1(\psi) = g\psi - s$. Write $F_N^1(\psi) = (1-\psi)(\eta - \lambda N\psi)$ for the drift under full effort. If $Ns > \eta h$, then $F_N^1 < 0$ on $[q, 1)$. The next lemma computes a player's surplus v over the safe payoff when all players wait below q, hold expected quality at q and use full effort above it, and her private gain G_1 from effort.

Lemma H.3 (The known resource)**.** *Suppose that $Ns > \eta h$. For $\psi \in (q, 1)$ let $T(\psi)$ be the date at which the full-effort path from ψ reaches q, and put $T(1) = \infty$. Let $v = 0$ on $[0, q]$ and, for $\psi \in (q, 1]$,*

$$v(\psi) = \int_0^{T(\psi)} e^{-r\tau} Q_\tau\, r\, m_1(\psi_\tau)\,\mathrm{d}\tau, \qquad Q_\tau = \exp\Big(-\lambda N \int_0^\tau \psi_\sigma\,\mathrm{d}\sigma\Big),$$

along the full-effort path (ψ_τ) from ψ. Then the following hold.

(a) The function v is continuous on $[0, 1]$, continuously differentiable on $[0, 1)$ with $v' = 0$ on $[0, q]$, and positive on $(q, 1]$.

(b) On $(q, 1)$, $F_N^1 v' = (r + \lambda N\psi)v - r\, m_1$.

(c) Let $G_1 = r\, m_1 - \lambda\psi[v + (1-\psi)v']$ on $[0, 1)$ and $G_1(1) = r\, m_1(1) - \lambda v(1)$. Then $G_1 < 0$ on $[0, q)$, $G_1(q) = 0$, $G_1(1) > 0$, and on $(q, 1)$

$$G_1 = \frac{r\, m_1(\eta - \lambda(N-1)\psi) - \lambda\psi(\eta + r)\, v}{\eta - \lambda N\psi}. \tag{H.5}$$

If $(N-1)s \geq \eta h$, then $G_1 > 0$ on $(q,1]$. If $(N-1)s < \eta h$, then $G_1 < 0$ on an interval $(q, q+\delta)$.

Proof. (a) and (b). Since F_N^1 is smooth, negative on $[q,1)$ and zero at one, T is finite and smooth on $(q,1)$, because the path crosses q with negative speed, and $T(\psi) \to 0$ as $\psi \downarrow q$. Before T, $m_1 > 0$, so $v > 0$ on $(q,1]$. Along the path, $e^{-r\tau}Q_\tau v(\psi_\tau) + \int_0^\tau e^{-r\sigma}Q_\sigma\, r\, m_1(\psi_\sigma)\,\mathrm{d}\sigma$ is constant, and its derivative at $\tau = 0$ gives (b). Since $m_1 \leq g(\psi - q)$ along the path and $T(\psi)$ is at most $\psi - q$ divided by the smallest value of $|F_N^1|$ on $[q,\psi]$, $v(\psi) = O\big((\psi-q)^2\big)$. By (b), $v'(\psi) \to 0$ as $\psi \downarrow q$, because the right side vanishes at q and $F_N^1(q) \neq 0$, so v is continuously differentiable on $[0,1)$. At $\psi = 1$ the path stays at one, so $v(1) = r(g-s)/(r+\lambda N)$, and v is continuous at one because the full-effort path depends continuously on its initial state and the integrand is bounded and discounted.

(c) Below q, $v = v' = 0$, so $G_1 = r\, m_1 < 0$, and $G_1(q) = 0$. Substituting v' from (b) gives (H.5), whose denominator is negative on $(q,1)$, and $G_1(1) = r(g-s)(r+\lambda(N-1))/(r+\lambda N) > 0$. If $\lambda(N-1)q \geq \eta$, then $\eta - \lambda(N-1)\psi < 0$ for $\psi > q$, so the numerator of (H.5) is negative and $G_1 > 0$. If $\lambda(N-1)q < \eta$, then, because $v = O\big((\psi-q)^2\big)$, the numerator equals $rg(\psi-q)\big(\eta - \lambda(N-1)q\big) + O\big((\psi-q)^2\big)$, which is positive near q. □

Let a_1 be the strategy of Proposition 7 on the edge, which waits below q, uses $\eta/(\lambda N q)$ at q and full effort above q, and let $\hat{u} = s + v$. When the other players can hold expected quality at q by themselves, this profile is an equilibrium of the game after a discovery.

Lemma H.4 (Equilibrium after a discovery)**.** *Suppose that $(N-1)s \geq \eta h$. On the edge $p = 1$, the profile in which every player uses a_1 is well posed against unilateral deviations, its profile path exists and is unique from every ψ, and it yields player i the payoff $\hat{u}(\psi)$, which no unilateral deviation exceeds. In particular, right after a discovery no deviation yields more than $\hat{u}(0) = s$.*

Proof. Since $\lambda(N-1)q \geq \eta$, the effort $\eta/(\lambda N q)$ is below one, and a_1 is nondecreasing, so the profile is well posed by Lemma H.1(b). From $\psi < q$ the profile path waits until ψ reaches q, and from $\psi \in (q,1)$ it uses full effort until ψ reaches q. At q the drift points toward q from both sides, so, as in the proof of Theorem 4, the unique solution stays at q with aggregate effort $\eta/(\lambda q) = N a_1(q)$. From $\psi = 1$ the state stays at one.

Payoffs before the next success. Fix $\psi_0 \in [0,1)$, a measurable schedule b of player i, its trajectory (ψ_t) and the opponents' effective effort A_{-i}, and put $A = b + A_{-i}$ and $Q_t = \exp(-\lambda \int_0^t A\psi)$, the probability of no success by t. We show that

$$\hat{u}(\psi_0) \geq \int_0^\infty e^{-rt} Q_t \big[r\big(s + b_t\, m_1(\psi_t)\big) + \lambda A_t \psi_t\, s\big]\,\mathrm{d}t, \tag{H.6}$$

with equality under the profile. The function $\Xi(t) = e^{-rt}Q_t\hat{u}(\psi_t) + \int_0^t e^{-r\sigma}Q_\sigma[r(s + b\, m_1) + \lambda A \psi s]\,\mathrm{d}\sigma$ is absolutely continuous, and almost everywhere $\Xi' = e^{-rt}Q_t\,\iota$ with

$$\iota = -rv + \eta(1-\psi)v' + b\, G_1 + A_{-i}(G_1 - r\, m_1).$$

Below q, $A_{-i} = 0$ and $v = v' = 0$, so $\iota = b\, r\, m_1 \leq 0$. Above q, $A_{-i} = N - 1$, and Lemma H.3(b) reads $rv = \eta(1-\psi)v' + NG_1 - (N-1)r\, m_1$, so $\iota = (b-1)G_1 \leq 0$. At q, $v = v' = m_1 = G_1 = 0$, so $\iota = 0$ whatever the effective efforts. Thus Ξ is nonincreasing, and since $\hat{u} \geq 0$, $\hat{u}(\psi_0) = \Xi(0)$ is at least the integral up to any date, which gives (H.6). Under the profile $\iota = 0$ almost everywhere and $e^{-rt}Q_t\hat{u}(\psi_t) \to 0$, which gives equality. From $\psi_0 = 1$ the state stays at one, $rv(1) = NG_1(1) - (N-1)r\, m_1(1)$, and the same argument gives $\iota = (b-1)G_1(1) \leq 0$.

Successive successes. Let $\bar{J}(\psi)$ be the supremum of player i's payoff over unilateral deviations from ψ, and let $B = \sup_\psi[\bar{J}(\psi) - \hat{u}(\psi)]$, which is finite because payoffs are bounded. Conditional on the deviator's private randomization, a deviation is a measurable schedule until the first success. Given the history up to the first success, the rest of the deviation is itself a unilateral deviation from $\psi = 0$, because the opponents' strategies and the law of motion do not depend on calendar time, and its conditional payoff is measurable in that history. Its continuation is therefore worth at most $\bar{J}(0) \leq s + B$. By (H.6),

$$J^i \leq \int_0^\infty e^{-rt}Q_t\big[r(s + b\, m_1) + \lambda A\psi(s + B)\big]\,\mathrm{d}t \leq \hat{u}(\psi_0) + \max\{B, 0\}\int_0^\infty e^{-rt}Q_t\lambda A_t\psi_t\,\mathrm{d}t.$$

The last integral equals $1 - r\int_0^\infty e^{-rt}Q_t\,\mathrm{d}t$, which is at most $\lambda N/(r + \lambda N)$ because $Q_t \geq e^{-\lambda Nt}$. Taking suprema gives $B \leq \lambda N \max\{B, 0\}/(r+\lambda N)$, so $B \leq 0$. The same computation with equalities shows that the profile's payoff J^* satisfies $J^*(\psi) - \hat{u}(\psi) = \gamma(\psi)\big(J^*(0) - s\big)$ with $0 \leq \gamma(\psi) < 1$, so $J^*(0) = s$ and $J^* = \hat{u}$. In particular $B = 0$. □

H.4 Before a discovery

Before a discovery, the reduced value of the strategy of Proposition 7 is zero on and below the myopic curve and, above it, the discounted myopic surplus collected under full effort until the state reaches the curve. Its private gain has a closed form. Put $\psi_N = \eta/(\lambda N)$. If $Ns > \eta h$, then $\psi_N < q$. Under full effort, (H.3) is a linear system with eigenvalues $-\eta$ and $-\lambda N$ and eigenvectors $(1, \psi_N)$ and $(1, 1)$, whose solution from (X, Y) is

$$(X_\tau, Y_\tau) = \frac{(X - Y)\,(1, \psi_N)\,e^{-\eta\tau} + (Y - \psi_N X)\,(1, 1)\,e^{-\lambda N\tau}}{1 - \psi_N}. \tag{H.7}$$

Let $W = 0$ on and below the myopic curve and, above it,

$$W(X, Y) = \int_0^T e^{-r\tau}\, r\, m(X_\tau, Y_\tau)\,\mathrm{d}\tau,$$

where T is the date at which the full-effort path (H.7) from (X, Y) meets the curve. Write $F_N(X, Y) = (-\lambda NY,\ \eta(X - Y) - \lambda NY)$ for the drift under full effort, and $f_1(X, Y) = -\lambda Y\,(1, 1)$ for the change in the drift per unit of own effort. The *private gain* is $G = r\, m + \nabla W \cdot f_1$, the counterpart of (6.2) when a discovery is worth nothing to a player.

Lemma H.5 (Value and gain before a discovery)**.** *Suppose that $Ns > \eta h$.*

(a) From every state above the myopic curve, the full-effort path stays above the curve until a finite date T, at which it crosses the curve with $p\psi$ strictly decreasing. The date T is smooth above the curve and tends to zero at the curve.

(b) The function W is continuously differentiable on $\{(X,Y) : X > 0,\ Y < X\}$, which contains every state with $0 < p < 1$ and $\psi < 1$. It vanishes on and below the curve, satisfies $0 \le W \le (g-s)X$, and above the curve $rW = r\,m + \nabla W \cdot F_N$.

(c) The private gain is negative below the curve, zero on it, and above it

$$G = r\Big[m - \lambda(g-s)\,Y\,\frac{1-e^{-(r+\lambda N)T}}{r+\lambda N}\Big]. \tag{H.8}$$

(d) If $(N-1)s \ge \eta h$, then $G > 0$ above the curve.

On the line $Y = X$ of full expected quality, which full effort leaves invariant, (b) to (d) hold for the restriction of W, with derivatives along the line.

Proof. (a) Under full effort ψ_N is a rest point of ψ, so a path that starts above it stays above it, and above the curve $\psi > q > \psi_N$. By (8.1), $\frac{d}{dt}(p\psi) = p\big[\eta(1-\psi) - \lambda N\psi(1-p\psi)\big]$, which is at most $p(1-\psi)(\eta - \lambda N\psi) < 0$ when $\psi_N < \psi < 1$ and equals $-\lambda N p(1-p) < 0$ when $\psi = 1$. While the path is above the curve, $\psi > q$, so $\dot p \le -\lambda N q\,p(1-p)$, and the belief would fall below q in finite time. Hence the path meets the curve at a finite date T, and $p\psi$ is strictly decreasing there. Since $m = g(p\psi - q)/(1-p)$, the implicit function theorem gives a smooth T, which vanishes on the curve.

(b) Above the curve, W is smooth because the flow (H.7) and T are. The derivative of W in any direction is the integral of $re^{-r\tau}$ times the derivative of m along the perturbed path, because the term from the change in T is multiplied by $m = 0$ at the exit point. As the state approaches the curve, $T \to 0$, so $W \to 0$ and $\nabla W \to 0$. Since W vanishes on and below the curve, it is continuously differentiable. Along the path, $p\psi$ falls and p falls, so $m = g(p\psi - q)/(1-p)$ falls, and $0 \le W \le m \le (g-s)X$. Finally $e^{-r\tau}W(X_\tau, Y_\tau) + \int_0^\tau e^{-r\sigma} r\,m\,\mathrm{d}\sigma$ is constant along the path, and its derivative at $\tau = 0$ gives $rW = r\,m + \nabla W \cdot F_N$.

(c) Below the curve $\nabla W = 0$, so $G = r\,m < 0$, and on the curve $m = 0$ and $\nabla W = 0$. Above the curve, by (H.7), perturbing the initial state by $\varepsilon(1,1)$ moves the full-effort path by $\varepsilon e^{-\lambda N\tau}(1,1)$ and m by $\varepsilon(g-s)e^{-\lambda N\tau}$. By the argument of (b), the derivative $W_{\mathbf{1}} = (1,1)\cdot\nabla W$ of W in the direction $(1,1)$ equals $r(g-s)\int_0^T e^{-(r+\lambda N)\tau}\,\mathrm{d}\tau$, and $G = r\,m - \lambda Y\,W_{\mathbf{1}}$ is (H.8).

(d) Along the full-effort path, the exit date from the current state is $T-\tau$, so $\dot W_{\mathbf{1}} = (r+\lambda N)W_{\mathbf{1}} - r(g-s)$, and $\dot m = g\eta(X-Y) - \lambda N(g-s)Y$ by (H.3). With $\kappa = \eta(X-Y)/Y = \eta(1-\psi)/\psi$ and $\lambda Y W_{\mathbf{1}} = r\,m - G$, differentiating $G = r\,m - \lambda Y W_{\mathbf{1}}$ gives

$$\frac{d}{d\tau}G = (r+\kappa)\,G + r\,\hat S_1, \qquad \hat S_1 = g\eta(X-Y) - \lambda(N-1)(g-s)Y - (r+\kappa)\,m. \tag{H.9}$$

Above the curve, $\kappa m = g\eta(X-Y) - s\eta(X-Y)(1+X)/Y$, with $s(1+X) < gY$, and $X - Y < Y(1-q)/q$

because $\psi > q$. Hence

$$\begin{aligned}\hat{S}_1 &= \frac{s\eta(X-Y)(1+X)}{Y} - \lambda(N-1)(g-s)Y - r\,m \\ &< g\eta(X-Y) - \lambda(N-1)(g-s)Y < \frac{(g-s)Y}{q}\big(\eta - \lambda(N-1)q\big) \le 0.\end{aligned}$$

Integrating (H.9) from the state to the exit point, where $G = 0$, gives

$$G = -r\int_0^T \exp\Big(-\int_0^\tau (r+\kappa)\,\mathrm{d}\sigma\Big)\hat{S}_1\,\mathrm{d}\tau > 0.$$

On the line $Y = X$, f_1 is tangent to the line, the full-effort path stays on it, $\kappa = 0$, and the same arguments apply to the restriction of W. □

In the units of the main text, $(1-p)\hat{S}_1 = \eta s(1-\psi)/\psi - \lambda(N-1)p\psi(g-s) - r(gp\psi - s)$, the counterpart of the source S_1 of (6.5) when a discovery is worth nothing to a player. The coefficient $r+\kappa$ in (H.9) exceeds the coefficient r of (6.6) because a unit of own effort shifts the state by $\lambda Y(1,1)$, and recovery raises Y along the path at the relative rate κ, whereas with observed failures a unit of effort shifts the log odds by the constant λ.

H.5 Proof of Proposition 7

Let a^* be the strategy of Proposition 7. On the curve, $\hat{A}_m(\psi) \le \hat{A}_m(q) = \eta/(\lambda q) \le N-1$, so a^* takes values in $[0,1]$, and on the edge it is a_1. Since $(N-1)s \ge \eta h$ implies $Ns > \eta h$, Lemmas H.3 and H.5 apply.

Well-posedness and the profile path. At a fixed value of the clock, p and $\psi = 1 - e^{c-x}$ increase with x, so a^* is nondecreasing in x, and Lemma H.1(a) gives well-posedness before a discovery from states with $\psi < 1$. At full expected quality Lemma 7.1 applies, and on the edge Lemma H.4. As in the proof of Theorem 4, the inclusion $\dot{x} \in -\lambda\psi N[a^*(x-), a^*(x+)]$ has a unique solution. At almost every date it spends on the curve, the derivative of $p\psi$ vanishes, and by the computation in the proof of Lemma H.5(a) with aggregate effort A, this means $\eta(1-\psi) = \lambda A\psi(1-q)$, that is $A = \hat{A}_m(\psi) = Na^*$. The solution is therefore the profile path, and it is a trajectory of every player's unilateral system with $u = \hat{A}_m/N$ and $A_{-i} = (N-1)\hat{A}_m/N \in [0, N-1]$.

Optimality before a discovery. Fix a state with $0 < p_0 < 1$, a deviation u, its trajectory $z_t = (X_t, Y_t)$ and the opponents' effective effort A_{-i}. The function $\Xi(t) = e^{-rt}W(z_t) + \int_0^t e^{-r\sigma} r\,u_\sigma m(z_\sigma)\,\mathrm{d}\sigma$ is absolutely continuous by Lemma H.5(b), and almost everywhere $\Xi' = e^{-rt}\iota$ with

$$\iota = -rW + \nabla W \cdot f_0 + A_{-i}\nabla W \cdot f_1 + u\,G, \qquad f_0(X,Y) = \big(0,\ \eta(X-Y)\big).$$

Below the curve $W = 0$, $\nabla W = 0$ and $A_{-i} = 0$, so $\iota = u\,r\,m \le 0$. Above it $A_{-i} = N-1$ and $rW = r\,m + \nabla W \cdot (f_0 + Nf_1)$, so $\iota = (u-1)G \le 0$ by Lemma H.5(d). On the curve $W = 0$, $\nabla W = 0$ and $m = 0$, so $\iota = 0$ whatever the effective efforts. If $\psi_0 = 1$, the trajectory stays on the line $Y = X$,

where $f_0 = 0$ and f_1 is tangent to the line, and the same computation applies to the restriction of W (Lemma H.5). Since $W \geq 0$ and $|m(z_t)| \leq (g+s)X_0 + s$ along the trajectory, because $X_t \leq X_0$, the reduced payoff of any deviation is at most $W(z_0)$. Under the profile $\iota = 0$ almost everywhere and $e^{-rt}W(z_t) \leq e^{-rt}(g-s)X_0 \to 0$, so the profile attains $W(z_0)$.

Optimality. By Lemma H.4, no deviation after a discovery yields more than s, so Lemma H.2 with $\bar{e} = 0$ gives $J^i - s \leq (1-p_0)W(z_0)$, with equality under the profile. At $p = 1$, Lemma H.4 applies. At $p = 0$ there are no opportunities and effort pays nothing, so waiting is optimal. Hence a^* is a symmetric Markov perfect equilibrium under the Filippov convention, with value $s + (1-p)W$ before a discovery and $\hat{u} = s + v$ on the edge, which gives (a) and (b).

Part (c). If $p_0 \leq q$, then $p\psi \leq q$ along the path and nobody experiments, so $p_t = p_0$. If $p_0 > q$, the path reaches the curve in finite time, by waiting, which raises ψ to $q/p_0 < 1$, or by full effort (Lemma H.5(a)), and then slides along it. On the curve $gY = s(1+X)$, and $X - Y = C_0 e^{-\eta t}$ for a constant $C_0 \geq 0$, so $X_t = (s + gC_0 e^{-\eta t})/(g-s) \to q/(1-q)$, that is $p_t \to q$. The probability of no discovery by t is $(1-p_0)/(1-p_t)$, which converges to $(1-p_0)/(1-\min\{p_0, q\})$.

The case $(N-1)s < \eta h < Ns$. Then $\lambda(N-1)q < \eta < \lambda N q$, so a_1 is a strategy, and $\hat{u} = s+v$ is its value on the edge, by the equality parts of the proof of Lemma H.4, which do not use $(N-1)s \geq \eta h$. By Lemma H.3(c), $G_1 < 0$ on some interval $(q, q+\delta)$, where the profile uses full effort. From ψ_0 in that interval, let player i wait for a short time τ and then follow the profile. With only the other $N-1$ players at full effort, ψ rises near q, because $\lambda(N-1)q < \eta$, so for τ small the state stays in the interval, and the function Ξ of the proof of Lemma H.4 satisfies $\Xi' = -e^{-rt}Q_t G_1 > 0$ on $(0, \tau)$. Player i's payoff is $\Xi(\tau) > \Xi(0) = \hat{u}(\psi_0)$, so the deviation is profitable. The same deviation from $\psi_0 = q$, a state that the profile path reaches after every success, is also profitable, because the other players' Filippov efforts at q lie in $[0, N-1]$ and cannot supply the aggregate effort $\eta/(\lambda q) > N-1$ that holds expected quality at q, so ψ enters $(q, q+\delta)$ at once. □

H.6 Numerical checks

The module `hidden.py` of the replication package computes every object of this appendix in closed form, and four checks support the proof. First, at the baseline the closed forms of v and W agree with numerical integration along full-effort paths to within 2×10^{-11} in relative terms, the gains (H.5) and (H.8) agree with finite differences of v and W to within 2×10^{-8}, and $(1-p)\hat{S}_1$ is negative at all 221,976 points of a 1,201 × 1,201 grid that lie above the curve. Second, we simulate the game with hidden quality, in which each path draws productivity, opportunities pay with the true quality and failures reset it unseen, with players following the equilibrium strategy from the belief 0.8 and the qualities 0, 0.4, 0.7 and 0.95. In 400,000 paths for each initial state, the share of paths with a discovery by date 40 lies between 0.599 and 0.601, against $(0.8-0.5)/(1-0.5) = 0.6$, with standard errors of 0.0008. Third, Table 4 reports the largest gain from a deviation on the edge $p = 1$ when players revise their effort every Δ, computed by dynamic programming. Fourth, it reports the same gain before a discovery, where the program runs backward in the clock c, which falls by $\eta\Delta$ in every step whatever the effort, so that only the log odds need a grid. When $(N-1)s \geq \eta h$, the gains fall

by about half or more each time Δ halves. When $(N-1)s < \eta h < Ns$, the gain on the edge stays near 6×10^{-3}, as Proposition 7 predicts.

Table 4: Largest gain from a deviation when failures are hidden

	Baseline		$N = 4$		$\eta = 0.5$		$\eta = 1.2$
Step Δ	after	before	after	before	after	before	after
0.02	2.4×10^{-3}	1.6×10^{-3}	3.2×10^{-3}	2.4×10^{-3}	7.6×10^{-4}	8.9×10^{-4}	5.5×10^{-3}
0.01	1.2×10^{-3}	8.8×10^{-4}	1.2×10^{-3}	9.1×10^{-4}	3.5×10^{-4}	4.4×10^{-4}	5.9×10^{-3}
0.005	5.3×10^{-4}	4.2×10^{-4}	5.1×10^{-4}	4.1×10^{-4}	1.9×10^{-5}	1.7×10^{-4}	6.1×10^{-3}

Notes: payoff units. Players revise their effort every Δ, and opponents use the equilibrium effort of Proposition 7 at the latest revision date. After: largest gain on the edge $p = 1$, over 2,002 expected qualities. Before: largest gain before a discovery, over states with beliefs in $[0.05, 0.95]$ and expected qualities in $[0, 0.99]$, with a grid step of Δ in the log odds. Other primitives at the baseline. $(N-1)s = \eta h$ at the baseline, $(N-1)s > \eta h$ for $N = 4$ and for $\eta = 0.5$, and $(N-1)s < \eta h < Ns$ for $\eta = 1.2$.

References

Axel Anderson, Lones Smith, and Andreas Park. Rushes in large timing games. *Econometrica*, 85 (3):871–913, 2017.

Elena Antoniadou, Christos Koulovatianos, and Leonard J. Mirman. Strategic exploitation of a common-property resource under uncertainty. *Journal of Environmental Economics and Management*, 65(1):28–39, 2013. doi: 10.1016/j.jeem.2012.05.005.

Simon Board and Moritz Meyer-ter Vehn. Experimentation in networks. *American Economic Review*, 114(9):2940–2980, 2024. doi: 10.1257/aer.20230233.

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

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

Alessandro Bonatti and Johannes Hörner. Learning to disagree in a game of experimentation. *Journal of Economic Theory*, 169:234–269, 2017.

Jorge Cortés. Discontinuous dynamical systems: A tutorial on solutions, nonsmooth analysis, and stability. *IEEE Control Systems Magazine*, 28(3):36–73, 2008.

Martin W. Cripps and Caroline D. Thomas. Strategic experimentation in queues. *Theoretical Economics*, 14(2):647–708, 2019.

Mark H. A. Davis. *Markov Models and Optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.

Engelbert J. Dockner, Steffen Jørgensen, Ngo Van Long, and Gerhard Sorger. *Differential Games in Economics and Management Science*. Cambridge University Press, Cambridge, 2000.

Ran Eilat, Zvika Neeman, and Eilon Solan. Opportunity hunters: A model of competitive sequential inspections. *American Economic Journal: Microeconomics*, 18(2):228–267, 2026. doi: 10.1257/mic.20240245.

Eric Fesselmeyer and Marc Santugini. Strategic exploitation of a common resource under environmental risk. *Journal of Economic Dynamics and Control*, 37(1):125–136, 2013. doi: 10.1016/j.jedc.2012.06.010.

Aleksei F. Filippov. *Differential Equations with Discontinuous Righthand Sides*, volume 18 of *Mathematics and Its Applications (Soviet Series)*. Kluwer Academic Publishers, Dordrecht, 1988.

Drew Fudenberg and Jean Tirole. Preemption and rent equalization in the adoption of new technology. *Review of Economic Studies*, 52(3):383–401, 1985.

H. Scott Gordon. The economic theory of a common-property resource: The fishery. *Journal of Political Economy*, 62(2):124–142, 1954.

Marina Halac, Navin Kartik, and Qingmin Liu. Contests for experimentation. *Journal of Political Economy*, 125(5):1523–1569, 2017.

Paul Heidhues, Sven Rady, and Philipp Strack. Strategic experimentation with private payoffs. *Journal of Economic Theory*, 159:531–551, 2015.

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.

Johannes Hörner, Nicolas Klein, and Sven Rady. Overcoming free-riding in bandit games. *Review of Economic Studies*, 89(4):1948–1992, 2022. doi: 10.1093/restud/rdab078.

Niko Jaakkola and Florian Wagener. Markov-perfect equilibria in differential games—with an application to climate policy. *The Review of Economic Studies*, 2026. doi: 10.1093/restud/rdag045. Advance article rdag045, published 14 May 2026.

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

Godfrey Keller and Sven Rady. Breakdowns. *Theoretical Economics*, 10(1):175–202, 2015. doi: 10.3982/TE1670.

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.

Kyungmin Kim and Frances Xu Lee. Information acquisition in a war of attrition. *American Economic Journal: Microeconomics*, 6(2):37–78, 2014. doi: 10.1257/mic.6.2.37.

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

Barbara Koeck, Magnus Lovén Wallerius, Robert Arlinghaus, and Jörgen I. Johnsson. Behavioural adjustment of fish to temporal variation in fishing pressure affects catchability: An experiment with angled trout. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(1):188–193, 2020. doi: 10.1139/cjfas-2019-0064.

David Levhari and Leonard J. Mirman. The great fish war: An example using a dynamic Cournot–Nash solution. *Bell Journal of Economics*, 11(1):322–334, 1980.

Eric Maskin and Jean Tirole. Markov perfect equilibrium: I. Observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

Andreas Park and Lones Smith. Caller number five and related timing games. *Theoretical Economics*, 3(2):231–256, 2008.

Juliusz Schauder. Der Fixpunktsatz in Funktionalräumen. *Studia Mathematica*, 2:171–180, 1930. doi: 10.4064/sm-2-1-171-180.

Caroline D. Thomas. Strategic experimentation with congestion. *American Economic Journal: Microeconomics*, 13(1):1–82, 2021. doi: 10.1257/mic.20170187.