LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

VOLKSWIRTSCHAFTLICHE FAKULTÄT

Godfrey Keller und Sven Rady:

Strategic Experimentation with Poisson Bandits

Munich Discussion Paper No. 2009-5

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Strategic Experimentation with Poisson Bandits*

Godfrey Keller† Sven Rady‡

This version: April 8, 2009

Abstract

We study a game of strategic experimentation with two-armed bandits where the risky arm distributes lump-sum payoffs according to a Poisson process. Its intensity is either high or low, and unknown to the players. We consider Markov perfect equilibria with beliefs as the state variable. As the belief process is piecewise deterministic, payoff functions solve differential-difference equations. There is no equilibrium where all players use cut-off strategies, and all equilibria exhibit an 'encouragement effect' relative to the single-agent optimum. We construct asymmetric equilibria in which players have symmetric continuation values at sufficiently optimistic beliefs yet take turns playing the risky arm before all experimentation stops. Owing to the encouragement effect, these equilibria Pareto dominate the unique symmetric one for sufficiently frequent turns. Rewarding the last experimenter with a higher continuation value increases the range of beliefs where players experiment, but may reduce average payoffs at more optimistic beliefs. Some equilibria exhibit an 'anticipation effect': as beliefs become more pessimistic, the continuation value of a single experimenter increases over some range because a lower belief means a shorter wait until another player takes over.

Keywords: Strategic Experimentation, Two-Armed Bandit, Poisson Process, Bayesian Learning, Piecewise Deterministic Process, Markov Perfect Equilibrium, Differential-Difference Equation.

JEL Classification Numbers: C73, D83, O32.

---

*Our thanks for very helpful discussions and suggestions are owed to Martin Cripps, Johannes Hörner and to seminar participants at the universities of Munich, Nürnberg, Oxford, Toulouse, Wisconsin–Madison, ETH Zurich, the 2007 Annual Meeting of the Society for Economic Dynamics in Prague, and the SFB/TR 15 Conference on Strategic Information Acquisition and Transmission (Munich, November 2008). We thank the Institut d'Economie Industrielle at the University of Toulouse, the Center for Economic Studies at the University of Munich, the Business and Public Policy Group at the Wharton School and the Social Science Research Center Berlin (WZB) for their hospitality. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 and GRK 801 and from the Bavarian Graduate Program in Economics is gratefully acknowledged.

†Department of Economics, University of Oxford, Manor Road Building, Oxford OX1 3UQ, UK.

‡Department of Economics, University of Munich, Kaulbachstr. 45, D-80539 Munich, Germany.

1 Introduction

When firms cooperate in a research joint venture, each faces a dynamic problem in which it can perform repeated costly experiments (that is, spend time, effort and money on the purported innovation) but also learn from the experimental observations of the others. Such a game of *strategic experimentation* arises in a variety of economic contexts; besides firms' research and development activities, consumer search or experimental consumption of a new product are prominent examples. Academic researchers pursuing a common research agenda or simply working on a joint paper are also effectively engaged in strategic experimentation.

In this paper we analyze a game of strategic experimentation where a finite number of players face identical two-armed bandit problems. There is a safe arm that offers a known and constant flow payoff and a risky arm whose lump-sum payoffs are driven by a Poisson process of unknown intensity. The risky arm can be either 'good' or 'bad': if it is good, the lump-sums arrive more frequently than if it is bad. While all risky arms are of the same type (all good, or all bad), lump-sums arrive independently across players. Each player is endowed with a stream of one unit of a perfectly divisible resource and, at each point in time, must decide how to split this resource between the two arms. Players' actions and outcomes are publicly observed, so there are perfect informational spillovers between players.

With Poisson bandits, news arrives in a 'lumpy' fashion. For concreteness, we focus on a situation where this news is good. Examples would be the occasional 'breakthrough' in research and development, a completed research paper in a longer-term research agenda, or one of a sequence of crucial proofs in a paper. Beliefs jump to a more optimistic level whenever a 'news event' or 'success' occurs, whereas they gradually become more pessimistic in between such events.

A single success on the risky arm does not fully reveal its type. This stands in marked contrast to the experimentation model of Keller, Rady and Cripps (2005). There, a good risky arm also generates lump-sum payoffs according to some Poisson process, but a bad risky arm never generates any payoffs, so the belief jumps all the way to certainty as soon as the first lump-sum arrives, irrespective of the belief held immediately before. In the present model, there is never certainty about the state of the risky arm, and the belief held immediately after a success varies with the belief held immediately before. As a consequence, when the players in our model use Markov strategies with the posterior belief as the state variable, their payoff functions solve first-order differential-difference equations. Despite this technical complication, these equations can be analyzed by elementary methods and admit closed-form solutions.

While all Markov perfect equilibria of the experimentation game are inefficient because of free-riding, we show that they always exhibit an 'encouragement effect': the presence of other players encourages at least one of them to continue experimenting with the risky arm at beliefs where a single agent would already have given up. This effect was first described by Bolton and Harris (1999) in a model where the risky arm yields a flow payoff with Brownian noise. Focusing on the symmetric MPE of their model, however, they obtain the encouragement effect for this particular equilibrium only. In contrast, we are able to establish this effect for *all* MPE of the Poisson model.

The unique symmetric MPE of the Poisson model shares the main features with its counterpart in the Bolton-Harris model. All players use the risky arm exclusively when they are sufficiently optimistic, the safe arm when they are sufficiently pessimistic, and both arms simultaneously at intermediate beliefs. Further, the acquisition of information is slowed down so severely near the lower bound of the intermediate range that the players' beliefs cannot reach this bound in finite time.

This strongly suggests that asymmetric equilibria where a last experimenter keeps the rate of information acquisition bounded away from zero at pessimistic beliefs ought to be more efficient than the symmetric one. Bolton and Harris (2000), who study the undiscounted limit of the Brownian model, and Keller, Rady and Cripps (2005) confirm this by constructing a variety of asymmetric MPE that dominate the symmetric one in terms of aggregate payoffs. However, they do so in environments without the encouragement effect. In fact, the existence and structure of asymmetric equilibria *with* an encouragement effect has remained an open question in the literature so far. The present paper fills this gap.

We show first that there is no MPE in which all players use cut-off strategies, i.e. use the risky arm exclusively when the probability they assign to the risky arm being good is above some cut-off, and the safe arm when it is below. In fact, the player who is supposed to use the least optimistic cut-off in a purported MPE in cut-off strategies always has an incentive to deviate to the safe action at the second least optimistic cut-off, where one of the other players is supposed to switch action.

We then construct, for an arbitrary number of players, asymmetric Markov perfect equilibria that generate a higher aggregate payoff than the symmetric MPE. They do so by combining behavior as in the symmetric equilibrium (at optimistic beliefs) with other behavior (at more pessimistic beliefs) where players take turns, one at a time, to play the risky arm exclusively while all others free-ride. As in Keller, Rady and Cripps (2005), the gain in aggregate payoffs stems from the fact that, owing to this alternation, the intensity of experimentation is bounded away from zero immediately above the belief where all experimentation stops. Because of the encouragement effect, this alternation can actually be performed in a way that leads to a Pareto improvement over the symmetric equilibrium. With sufficiently frequent switching between the roles of experimenter and free-rider, *every* player's payoff function closely approaches the average payoff function, making even the last experimenter better off than in the symmetric MPE.

While these equilibria require players to use interior allocations of their resource over some intermediate range of beliefs, we also construct what we call 'simple' equilibria, that is, MPE where at each belief, each player allocates his entire resource to one or other of the two arms. We do so for two players and under the assumption that the frequency of lump-sums on a bad risky arm is sufficiently low, which implies that after the arrival of a lump-sum payoff, all players revert to exclusive use of their risky arms. This allows us in particular to study the robustness of the simple equilibria in Keller, Rady and Cripps (2005) to the introduction of breakthroughs that are not fully revealing.

In a last step, we give examples of equilibria where the two players have asymmetric continuation values after a success on a risky arm; all MPE constructed up to that point actually have symmetric post-success continuation values. Introducing these asymmetries allows us to reward the last experimenter. We find that this raises the average payoff at relatively pessimistic beliefs, but lowers them at optimistic beliefs. The local increase in average payoffs is once more a consequence of the encouragement effect; in Keller, Rady and Cripps (2005), where the encouragement effect is not present, making the players' equilibrium payoffs less symmetric at optimistic beliefs uniformly lowers the average payoff.

We also find a new 'anticipation effect' in the alternation phase of the examples we calculate: for some parameter values, the value function of a lone experimenter is *decreasing* in the current belief over some range. The intuition for this effect is that, conditional on no impending success on his own risky arm, the player will soon be able to enjoy a free-ride, and the lower the current belief, the sooner this time will come.

The Poisson model is a natural analog in continuous time of the two-outcome bandit model in Rothschild (1974), the first paper to use the bandit framework in economics; see Bergemann and Välimäki (2008) for a survey of the ensuing literature. Through its focus on bandit learning as a canonical model of strategic experimentation in teams, our paper is most closely related to Bolton and Harris (1999, 2000) and Keller, Rady and Cripps (2005), sharing with them the assumption that the players face risky arms of a common type. Klein and Rady (2008) and Klein (2009), by contrast, consider two players who face risky arms of opposite types, one good and one bad, with uncertainty about who has the good arm. Strulovici (2008) studies majority voting in a collective decision problem where the type of the risky arm also varies across individuals.

Our paper is further related to a strand of the industrial organization literature that studies R&D investment games under learning. Malueg and Tsutsui (1997) investigate a model of a patent race with learning where the arrival time of an innovation is exponentially distributed given the stock of knowledge, implying the same deterministic belief revision prior to the innovation as our model exhibits in between lump-sums. Building on the exponential bandit framework of Keller, Rady and Cripps (2005), Besanko and Wu (2008) study the effects of post-innovation market structure on cooperative and competitive R&D investments, respectively. Décamps and Mariotti (2004), Hopenhayn and Squintani (2008) and Moscarini and Squintani (2007) all analyze models where news arrives in the form of the increments of a (compound) Poisson process; as they consider stopping games with private information, however, the resulting strategic interactions are very different from that in our model.

The remainder of the paper is organized as follows. Section 2 sets up the Poisson bandit model. Section 3 establishes the efficient benchmark. Section 4 introduces the strategic problem and establishes the encouragement effect. Section 5 presents the unique symmetric MPE. Section 6 proves the impossibility of cut-off equilibria and constructs asymmetric equilibria. Section 7 studies simple equilibria for two players. Section 8 contains concluding remarks. Some of the proofs are relegated to the Appendix.

2 Poisson Bandits

The set-up of the model is similar to that of Keller, Rady and Cripps (2005), the principal difference being that here a bad risky arm yields positive payoffs (as opposed to zero), which means that a success does not reveal the risky arm to be good. For mathematical details on Poisson bandits, see Presman (1990) or Presman and Sonin (1990); for the optimal control of piecewise deterministic processes more broadly, see Davis (1993).

Time $t \in [0, \infty)$ is continuous, and the discount rate is $r > 0$. There are $N \geq 1$ players, each of them endowed with one unit of a perfectly divisible resource per unit of time, and each facing a two-armed bandit problem. Lump-sums rewards on the risky arm R are independent draws from a time-invariant distribution on $\mathbb{R}_{++}$ with a known mean h. If a player allocates the fraction $k_t \in [0, 1]$ of her resource to R over an interval of time $[t, t + dt[$, and consequently the fraction $1 - k_t$ to the safe arm S, then she receives the expected payoff $(1 - k_t)s\,dt$ from S, where $s > 0$ is a constant known to all players. The probability that she receives a lump-sum payoff from R at some point in the interval is $k_t\lambda_\theta\,dt$, where $\theta = 1$ if R is good, $\theta = 0$ if R is bad, and $\lambda_1 > \lambda_0 > 0$ are constants known to all players. Therefore, the overall expected payoff increment conditional on θ is $[(1 - k_t)s + k_t\lambda_\theta h]\,dt$. We assume that $\lambda_0 h < s < \lambda_1 h$, so each player prefers R to S if R is good, and prefers S to R if R is bad.

However, players do not know whether the risky arm is good or bad; they start with a common prior belief about θ. Thereafter, all players observe each other's actions and outcomes, so they hold common posterior beliefs throughout time. With p_t denoting the subjective probability at time t that players assign to the risky arm being good, a player's expected payoff increment conditional on all available information is $[(1 - k_t)s + k_t\lambda(p_t)h]\,dt$ with

$$\lambda(p) = p\lambda_1 + (1 - p)\lambda_0.$$

Given a player's actions $\{k_t\}_{t\geq 0}$ such that k_t is measurable with respect to the information available at time t, her total expected discounted payoff, expressed in per-period units, is

$$\mathrm{E}\left[\int_0^\infty r\,e^{-r\,t}\;[(1 - k_t)s + k_t\lambda(p_t)h]\;dt\right],$$

where the expectation is over the stochastic processes $\{k_t\}$ and $\{p_t\}$. We note that a player's payoff depends on others' actions only through their effect on the evolution of beliefs, which constitute a natural state variable.

To derive the law of motion of beliefs, suppose that over the interval of time $[t, t+\Delta t[$ player n allocates the constant fraction $k_{n,t}$ of her resource to her risky arm. The sum $K_t = \sum_{n=1}^N k_{n,t}$ measures how much of the overall resource is allocated to risky arms; we will call this number the *intensity of experimentation*. Conditional on the type of the risky arm, the arrival of lump-sums is independent across players. If the risky arms are good, the probability of none of the players receiving a lump-sum payoff is $e^{-K_t\lambda_1\Delta t}$, and if they are bad, this probability is $e^{-K_t\lambda_0\Delta t}$. Therefore, given no lump-sum payoff arriving in $[t, t + \Delta t[$, the belief at the end of that time period is

$$p_{t+\Delta t} = \frac{p_t\,e^{-K_t\lambda_1\Delta t}}{(1 - p_t)\,e^{-K_t\lambda_0\Delta t} + p_t\,e^{-K_t\lambda_1\Delta t}}$$

by Bayes' rule. As long as no lump-sum arrives, the belief thus evolves smoothly with infinitesimal increment $dp_t = -K_t\Delta\lambda\,p_t(1 - p_t)\,dt$ where $\Delta\lambda = \lambda_1 - \lambda_0$. However, if *any* of the players receives a lump-sum at time t, the belief jumps up from p_{t-} (the limit of beliefs held before the arrival of the lump-sum) to

$$p_t = \lim_{\Delta t \downarrow 0} \frac{p_{t-}\left[1 - e^{-K_t \lambda_1 \Delta t}\right]}{(1-p_{t-})\left[1 - e^{-K_t \lambda_0 \Delta t}\right] + p_{t-}\left[1 - e^{-K_t \lambda_1 \Delta t}\right]} = \frac{\lambda_1 p_{t-}}{\lambda(p_{t-})} \,,$$

which is independent of the intensity of experimentation. We write

$$j(p) = \frac{\lambda_1 p}{\lambda(p)}$$

for the function that describes beliefs after a success on a risky arm.

We restrict players to Markovian strategies $k_n : [0,1] \to [0,1]$ with the left limit belief p_{t-} as the state variable, so that the action player n takes at time t is $k_n(p_{t-})$.[1] We impose the following restrictions on these strategies: (i) k_n is left-continuous; (ii) there is a finite partition of $[0,1]$ into intervals of positive length on each of which k_n is Lipschitz-continuous. By standard results, each profile $(k_1, k_2, \ldots, k_N)$ of such strategies induces a well-defined law of motion for players' common beliefs and well-defined payoff functions. A *simple* strategy is one that takes values in $\{0,1\}$ only, meaning that the player uses one arm exclusively at any given point in time. Finally, a strategy k_n is a *cut-off* strategy if there is a belief $\hat{p}$ such that $k_n(p) = 1$ for all $p > \hat{p}$, and $k_n(p) = 0$ otherwise.

As a benchmark, a myopic agent would simply weigh the short-run payoff from playing the safe arm, s, against what he expects from playing the risky arm, $\lambda(p)h$. So he would use the cut-off belief

$$p^m = \frac{s - \lambda_0 h}{\Delta\lambda\, h},$$

playing R for $p > p^m$, and S for $p \leq p^m$.

3 Joint Maximization of Average Payoffs

Consider N players jointly maximizing their *average* expected payoff. By the same arguments as in Keller, Rady and Cripps (2005), the value function for the cooperative, expressed as average payoff per agent, satisfies the Bellman equation

$$u(p) = s + \max_{K \in [0,N]} K \left\{ b(p,u) - c(p)/N \right\},$$

[1] By definition, $p_{0-} = p_0$. Note that $p_{t-} = p_t$ at almost all t. Working with p_{t-} instead of p_t merely enforces the informational restriction that the action taken at time t cannot be conditioned on the arrival of a lump-sum at that time.

where K is the intensity of experimentation,

$$c(p) = s - \lambda(p)h$$

is the opportunity cost of playing R, and

$$b(p,u) = [\lambda(p)\,(u(j(p)) - u(p)) - \Delta\lambda\, p(1-p)u'(p)]/r$$

is the expected benefit of playing R. The latter has two parts: a discrete improvement in the overall payoff after a success, and a marginal decrease otherwise.[2]

If the shared opportunity cost of playing R exceeds the full expected benefit, the optimal choice is $K = 0$ (all agents use S exclusively), and $u(p) = s$. Otherwise, $K = N$ is optimal (all agents use R exclusively), and u satisfies the first-order ordinary differential-difference equation (henceforth ODDE)

$$\Delta\lambda\, p(1-p)u'(p) - \lambda(p)[u(j(p)) - u(p)] + \tfrac{r}{N}u(p) = \tfrac{r}{N}\lambda(p)h. \tag{1}$$

A particular solution to this equation is $u(p) = \lambda(p)h$, the expected per capita payoff from all agents using the risky arm forever.

The option value of being able to change to the safe arm is then captured by the solution to the homogeneous equation, for which we try $u_0(p) = (1-p)\Omega(p)^\mu$ for some $\mu > 0$ to be determined, where

$$\Omega(p) = \frac{1-p}{p}$$

is the odds ratio.[3] Now,

$$u_0'(p) = -\frac{\mu + p}{p(1-p)}\,u_0(p) \quad \text{and} \quad u_0(j(p)) = \frac{\lambda_0}{\lambda(p)}\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu} u_0(p).$$

Inserting these into the homogeneous equation and simplifying leads to the requirement

[2]Infinitesimal changes of the belief are always downward, so it is in fact the left-hand derivative of the value function that matters here. This observation will turn out to be of relevance in asymmetric equilibria of the strategic experimentation game.

[3]This guess can be obtained by 'extrapolation' from the limiting case $\lambda_0 = 0$ studied in Keller, Rady and Cripps (2005). In this case, $j(p) = 1$ and $u(j(p)) = \lambda_1 h$, so (1) becomes a linear differential equation; the above function u_0 is easily seen to solve the corresponding homogeneous equation for $\mu = r/(N\lambda_1)$. A more systematic approach relies on a change of the independent variable from p to $\ln\Omega(p)$. This transforms (1) into a linear ODDE with constant delay to which results from Bellman and Cooke (1963) can be applied.

that

$$\frac{r}{N} + \lambda_0 - \mu \Delta\lambda = \lambda_0 \left(\frac{\lambda_0}{\lambda_1}\right)^{\mu}. \tag{2}$$

As a function of μ, the left-hand side of (2) is a negatively sloped straight line which cuts the vertical axis at $\frac{r}{N} + \lambda_0$. The right-hand side is a decreasing exponential function which tends to 0 as $\mu \to +\infty$, tends to ∞ as $\mu \to -\infty$, and cuts the vertical axis at λ_0. Thus the above equation in μ has two solutions, one positive and one negative; we write μ_N for the positive solution, which obviously lies between $\frac{r}{N\Delta\lambda}$ (the value of μ where the left-hand side of (2) equals λ_0) and $\frac{r}{N\Delta\lambda} + \frac{\lambda_0}{\Delta\lambda}$ (the value of μ where it equals 0). As the left-hand side of (2) rises with $\frac{r}{N}$, we also see that μ_N is increasing in the discount rate and decreasing in the number of agents.

The solution to the ODDE for $K = N$ is thus

$$V_N(p) = \lambda(p)h + C\,(1-p)\,\Omega(p)^{\mu_N}, \tag{3}$$

where C is a constant of integration.[4]

Proposition 1 (Cooperative solution) *In the N-agent cooperative problem, there is a cut-off belief $p_N^* < p^m$ given by*

$$p_N^* = \frac{\mu_N(s - \lambda_0 h)}{(\mu_N + 1)(\lambda_1 h - s) + \mu_N(s - \lambda_0 h)} \tag{4}$$

such that below the cut-off it is optimal for all to play S exclusively and above it is optimal for all to play R exclusively. The value function V_N^* for the N-agent cooperative is given by

$$V_N^*(p) = \lambda(p)h + c(p_N^*)\left(\frac{1-p}{1-p_N^*}\right)\left(\frac{\Omega(p)}{\Omega(p_N^*)}\right)^{\mu_N} \tag{5}$$

when $p > p_N^*$, and $V_N^*(p) = s$ otherwise.

PROOF: The expression for p_N^* and the constant of integration in (5) are obtained by imposing $V_N^*(p_N^*) = s$ (value matching) and $(V_N^*)'(p_N^*) = 0$ (smooth pasting). Then, $b(p, V_N^*)$ falls short of $c(p)/N$ to the left of p_N^*, coincides with it at p_N^*, and exceeds it to the right of p_N^*. So V_N^* solves the Bellman equation, with the maximum being achieved at the intensity of experimentation stated in the proposition. ■

[4]The planner's solution in Bolton and Harris (1999) has the same structure. Only the expression for the expected current payoff from a risky arm and the exponent of the odds ratio differ across set-ups. Cohen and Solan (2008) show that this continues to be true when the risky arm generates payoffs according to a Lévy process (that is, a continuous-time process with stationary independent increments) with a binary prior on its characteristics.

The above proposition determines the *efficient* strategies. As in Bolton and Harris (1999) and Keller, Rady and Cripps (2005), it is efficient to use a common cut-off strategy; the cut-off increases in s and μ_N (and hence in $\frac{r}{N}$). The efficient intensity of experimentation exhibits a bang-bang feature, being maximal when the current belief is above p^*_N, and minimal when it is below.

4 The Strategic Problem

From now on, we assume that there are $N \geq 2$ players acting non-cooperatively. Our solution concept is Markov perfect equilibrium with the common belief as the state variable.

With $K_{\neg n}(p) = \sum_{\ell \neq n} k_\ell(p)$, and $b(p, u_n)$ and $c(p)$ as defined in Section 3 above, player n's payoff function u_n is continuous, piecewise differentiable, and satisfies

$$u_n(p) = s + K_{\neg n}(p)\, b(p, u_n) + k_n(p)\, \{b(p, u_n) - c(p)\}$$

on $[0, 1]$, with the second term on the right-hand side measuring the benefit of the information generated by the other players.

A strategy k^*_n for player n is a best response against his opponents' strategies if and only if the resulting payoff function u_n solves the Bellman equation

$$u_n(p) = s + K_{\neg n}(p)\, b(p, u_n) + \max_{k_n \in [0,1]} k_n\, \{b(p, u_n) - c(p)\}$$

on $[0, 1]$, and $k^*_n(p)$ achieves the maximum on the right-hand side at each belief p. It is straightforward to show that if player n plays a best response, his benefit of experimentation $b(p, u_n)$ is non-negative at all beliefs, and his payoff function u_n is non-decreasing in the other players' experimentation schedule, $K_{\neg n}$. Standard results further imply that a best-response payoff function u_n is once continuously differentiable at any point of continuity of $K_{\neg n}$.

At the boundaries of the unit interval, the obvious optimal actions are $k^*_n(0) = 0$ and $k^*_n(1) = 1$, which implies $u_n(0) = s$ and $u_n(1) = \lambda_1 h$ for the player's payoff function. More generally, player n's best response is obtained by comparing the opportunity cost of playing R with the expected *private* benefit. If $c(p) > b(p, u_n)$, then $k^*_n(p) = 0$, and the Bellman equation implies $u_n(p) = s + K_{\neg n}(p)\, b(p, u_n) < s + K_{\neg n}(p)\, c(p)$. If $c(p) = b(p, u_n)$, then $k^*_n(p)$ is arbitrary in $[0, 1]$, and $u_n(p) = s + K_{\neg n}(p)\, c(p)$. Finally, if $c(p) < b(p, u_n)$, then $k^*_n(p) = 1$, and $u_n(p) = s + (K_{\neg n}(p) + 1)\, b(p, u_n) - c(p) >$

$s + K_{\neg n}(p)\, c(p)$. Thus, exactly as in Keller, Rady and Cripps (2005), player n's best response to a given intensity of experimentation $K_{\neg n}$ depends on whether in the (p, u)-plane, the point $(p, u_n(p))$ lies below, on, or above the line

$$\mathcal{D}_{K_{\neg n}} = \{(p, u) \in [0, 1] \times I\!\!R_+ \colon\; u = s + K_{\neg n}\, c(p)\}.$$

For $K_{\neg n} > 0$ this is a downward sloping diagonal that cuts the safe payoff line $u = s$ at the myopic cut-off p^m; for $K_{\neg n} = 0$, it coincides with the safe payoff line.

The following two observations also carry over *verbatim* from Keller, Rady and Cripps (2005). First, no profile of Markov strategies can generate an average payoff that exceeds V_N^*, and the payoff of a player using a best response to her opponents' strategies cannot fall below V_1^*. The upper bound follows immediately from the fact that the cooperative solution maximizes the average payoff. The intuition for the lower bound is that an agent can only benefit from the information generated by others. Second, all Markov perfect equilibria are inefficient. Along the efficient experimentation path, the benefit of an additional experiment tends to $1/N$ of its opportunity cost as p approaches p_N^*. A self-interested player compares the benefit of an additional experiment with the full opportunity cost and so has an incentive to deviate from the efficient path by using S instead of R.

It is obvious that in any Markov perfect equilibrium, at least one player must be using the risky arm at any belief above the single-agent optimum p_1^*. The interesting question is whether experimentation continues *below* p_1^*, i.e. whether there is an encouragement effect. This effect rests on two conditions: the experimentation by any 'pioneer' contemplating the use of the risky arm below p_1^* must increase the likelihood that other players will return to the risky arm in the future, and these future actions must be valuable to the pioneer. In Keller, Rady and Cripps (2005), the encouragement effect is absent since the first success on the risky arm is fully revealing and so the second condition fails.

Bolton and Harris (1999) show that the encouragement effect is present in the symmetric Markov perfect equilibrium of their model. The next result shows that *all* MPE of our model exhibit the encouragement effect.

Proposition 2 (Encouragement effect) *In any Markov perfect equilibrium, at least one player experiments at some beliefs below p_1^*.*

PROOF: Suppose to the contrary that all players play S at all beliefs $p \leq p_1^*$. Then each player's payoff function satisfies $u_n(p_1^*) = s$ with the left-hand derivative

$u_n'(p_1^*-) = 0$. For S to be optimal we must have $b(p_1^*, u_n) \leq c(p_1^*) = b(p_1^*, V_1^*)$, and hence $u_n(j(p_1^*)) \leq V_1^*(j(p_1^*))$, which must in fact hold as an equality. Thus, the difference $u_n - V_1^*$ assumes its minimum (of 0) at $j(p_1^*)$, which implies $u_n'(j(p_1^*)-) \leq (V_1^*)'(j(p_1^*)-)$. As $u_n(j^2(p_1^*)) \geq V_1^*(j^2(p_1^*))$, this implies $b(j(p_1^*), u_n) \geq b(j(p_1^*), V_1^*)$ and hence $b(j(p_1^*), u_n) > c(j(p_1^*))$. So all players must use R at the belief $j(p_1^*)$. By the ODDE for V_1^* and the explicit solution in Proposition 1, we have $b(j(p_1^*), V_1^*) = V_1^*(j(p_1^*)) - s + c(j(p_1^*)) = V_1^*(j(p_1^*)) - \lambda(j(p_1^*))h > 0$. Each player's Bellman equation now yields

$$
\begin{aligned}
u_n(j(p_1^*)) &= s + N\, b(j(p_1^*), u_n) - c(j(p_1^*)) \\
&\geq s + N\, b(j(p_1^*), V_1^*) - c(j(p_1^*)) \\
&> s + b(j(p_1^*), V_1^*) - c(j(p_1^*)) \\
&= V_1^*(j(p_1^*)),
\end{aligned}
$$

which contradicts the equality $u_n(j(p_1^*)) = V_1^*(j(p_1^*))$ derived earlier. ■

The idea behind the proof is that the only way that all experimentation could stop at p_1^* is for the 'jump-benefit' to be the same for each of the N players as for a lone agent, given the same opportunity cost and the same 'slide-disbenefit'; but this would imply that u_n and V_1^* matched in value not only at p_1^* but also at $j(p_1^*)$. This is not possible, since at $j(p_1^*)$ the benefit of a further jump up is no less and the disbenefit of a slide down is no worse for player n than for a lone agent, and if a lone agent has an incentive to experiment then so do each of the N players, the positive externality resulting in a higher value at $j(p_1^*)$.

We now turn to a detailed investigation of Markov perfect equilibria.

5 Symmetric Equilibrium

A symmetric Markov perfect equilibrium admits three possible cases at any given belief. First, when all players play S exclusively, the common payoff is $u(p) = s$. Second, when all players play R exclusively, the common payoff function u satisfies (1), hence is of the form V_N given in (3). Third, when all players divide the resource between S and R, the indifference condition $b(p, u) = c(p)$ implies that the common payoff function solves the ODDE

$$
\Delta\lambda\, p(1-p)u'(p) - \lambda(p)[u(j(p)) - u(p)] = r\lambda(p)h - rs. \tag{6}
$$

In the (p, u)-plane, the region where all players use the risky arm exclusively and the region where they use both arms simultaneously are separated by the diagonal $\mathcal{D}_{N-1}$. Given the post-jump value $u(j(p))$, we have smooth pasting of the solutions to (1) and (6) along $\mathcal{D}_{N-1}$. Smooth pasting also occurs at the boundary of the region where all players use S exclusively with the region where they use both arms. In other words, u must be of class C^1. To see this, suppose we had a symmetric equilibrium with a payoff function that hits the level s at the belief $\tilde{p}$ with slope $u'(\tilde{p}+) > 0$. Then, at beliefs immediately to the right of $\tilde{p}$, we would have $b(p, u) = c(p)$ or

$$\lambda(p)[u(j(p)) - u(p)]/r = c(p) + \Delta\lambda\, p(1-p)u'(p)/r$$

implying

$$\lambda(\tilde{p})[u(j(\tilde{p})) - s]/r = c(\tilde{p}) + \Delta\lambda\, \tilde{p}(1-\tilde{p})u'(\tilde{p}+)/r > c(\tilde{p})$$

by continuity. Immediately to the left of $\tilde{p}$, continuity of $u(j(p))$ and the fact that $u'(p) = 0$ would then imply $b(p, u) = \lambda(p)[u(j(p)) - s]/r > c(p)$, so there would be an incentive to deviate from S to R.

Proposition 3 (Symmetric equilibrium) *The N-player experimentation game has a unique symmetric Markov perfect equilibrium with the common posterior belief as the state variable. The corresponding payoff function is the unique function $W_N\colon [0,1] \to [s, \lambda_1 h]$ of class C^1 with the following properties: $W_N(p) = s$ on an interval $[0, \tilde{p}_N]$ with $p_N^* < \tilde{p}_N < p_1^*$; $W_N(p) > s$ on $]\tilde{p}_N, 1]$; W_N solves (6) on an interval $]\tilde{p}_N, p_N^\dagger[$ with $\tilde{p}_N < p_N^\dagger < p^m$, and (1) on $]p_N^\dagger, 1[$. The players' common equilibrium strategy is continuous in the posterior belief and satisfies $k(p) = 0$ for $p \le \tilde{p}_N$,*

$$k(p) = \frac{1}{N-1}\frac{W_N(p) - s}{c(p)} \in\ ]0, 1[$$

for $\tilde{p}_N < p < p_N^\dagger$, and $k(p) = 1$ for $p \ge p_N^\dagger$. W_N is increasing on $[\tilde{p}_N, 1]$, and k on $[\tilde{p}_N, p_N^\dagger]$.

Proof: We first show that there is at most one symmetric MPE. Suppose that we have two symmetric equilibria with different payoff functions u and $\hat{u}$, respectively, both of which must be of class C^1. Let $u - \hat{u}$ assume a negative global minimum at the belief p, which by necessity must lie in the open unit interval. At this belief, $u'(p) = \hat{u}'(p)$ and $u(j(p)) - \hat{u}(j(p)) \ge u(p) - \hat{u}(p)$, so $b(p, u) \ge b(p, \hat{u})$. We cannot have both $u(p)$ and $\hat{u}(p)$ above $\mathcal{D}_{N-1}$ since in this region both u and $\hat{u}$ are of the form (3) and the difference $u - \hat{u}$ is increasing to the right of $\mathcal{D}_{N-1}$. Further, if $\hat{u}(p)$ is above $\mathcal{D}_{N-1}$ and $u(p)$ is on or below, then $b(p, \hat{u}) > c(p) = b(p, u)$ in contradiction to what we derived before. Consequently, we must have both $\hat{u}(p)$ and $u(p)$ on or below $\mathcal{D}_{N-1}$, so $b(p,\hat{u}) = c(p) = b(p,u)$. This in turn yields $u(j(p)) - \hat{u}(j(p)) = u(p) - \hat{u}(p)$, so the difference $u - \hat{u}$ is also at its minimum at the belief $j(p)$. Iterating the argument until we get to the right of p^m (and hence to the right of $\mathcal{D}_{N-1}$), we obtain another contradiction, which proves that $u \geq \hat{u}$. By the same arguments, $\hat{u} - u$ cannot assume a negative global minimum either, and so $u = \hat{u}$.

Next, we sketch the construction of the symmetric equilibrium; for details, see the Appendix. Varying the point of intersection with the diagonal $\mathcal{D}_{N-1}$, one first constructs a family of candidate value functions that solve the ODDE (1) (N players using R exclusively) above $\mathcal{D}_{N-1}$, and the ODDE (6) (indifference between R and S) below. Using an intermediate-value argument, one then establishes the existence of one such function that reaches the level s with zero slope as we move down from $p = p^m$ to lower beliefs. This function is easily seen to solve each player's Bellman equation. Finally, the identity $u_n(p) = s + K_{\neg n}(p)\, c(p)$ uniquely determines the common intensity of experimentation in the range of beliefs where the value function lies below $\mathcal{D}_{N-1}$ but above the level s. ■

We can represent the equilibrium payoff function W_N in closed form up to some constants of integration that are implicitly determined by $p_N^\dagger$.

Corollary 1 *Define intervals* $J_0 = [p_N^\dagger, 1]$ *and* $J_i = [j^{-i}(p_N^\dagger), j^{-(i-1)}(p_N^\dagger)[$ *for* $i = 1, 2, \ldots$. *If* $\mu_N \neq \lambda_0/\Delta\lambda$,[5] *then*

$$\begin{aligned} W_N(p) &= \lambda(p)h + i\left[\tfrac{r}{\lambda_1}(\lambda_1 h - s)p - \tfrac{r}{\lambda_0}(s - \lambda_0 h)(1-p)\right] \\ &\quad + C^{(0)} \left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu_N}}{\lambda_0 - \mu_N \Delta\lambda}\right)^i (1-p)\,\Omega(p)^{\mu_N} \\ &\quad + \sum_{\eta=0}^{i-1} \frac{C^{(i-\eta)}}{\eta!} \left(-\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\lambda_0/\Delta\lambda}}{\Delta\lambda} \ln\left[(\lambda_0/\lambda_1)^{\eta-1}\,\Omega(p)\right]\right)^{\eta} (1-p)\,\Omega(p)^{\lambda_0/\Delta\lambda} \end{aligned}$$

on $J_i \cap \{p : W_N(p) > s\}$ *for some constants* $C^{(i-\eta)}$ $(\eta = 0, \ldots i-1)$, *chosen to ensure continuity of* W_N. *The constant* $C^{(0)}$ *ensuring that* $(p_N^\dagger, W_N(p_N^\dagger)) \in \mathcal{D}_{N-1}$ *is given by*

$$C^{(0)} = N\, c(p_N^\dagger)\,(1 - p_N^\dagger)^{-1}\,\Omega(p_N^\dagger)^{-\mu_N}.$$

Proof: See the Appendix. The proof shows how the constants $C^{(i)}$ can be calculated recursively given $C^{(0)}$. ■

[5] The proof makes it obvious how one has to modify this result in the knife-edge case where $\mu_N = \lambda_0/\Delta\lambda$.

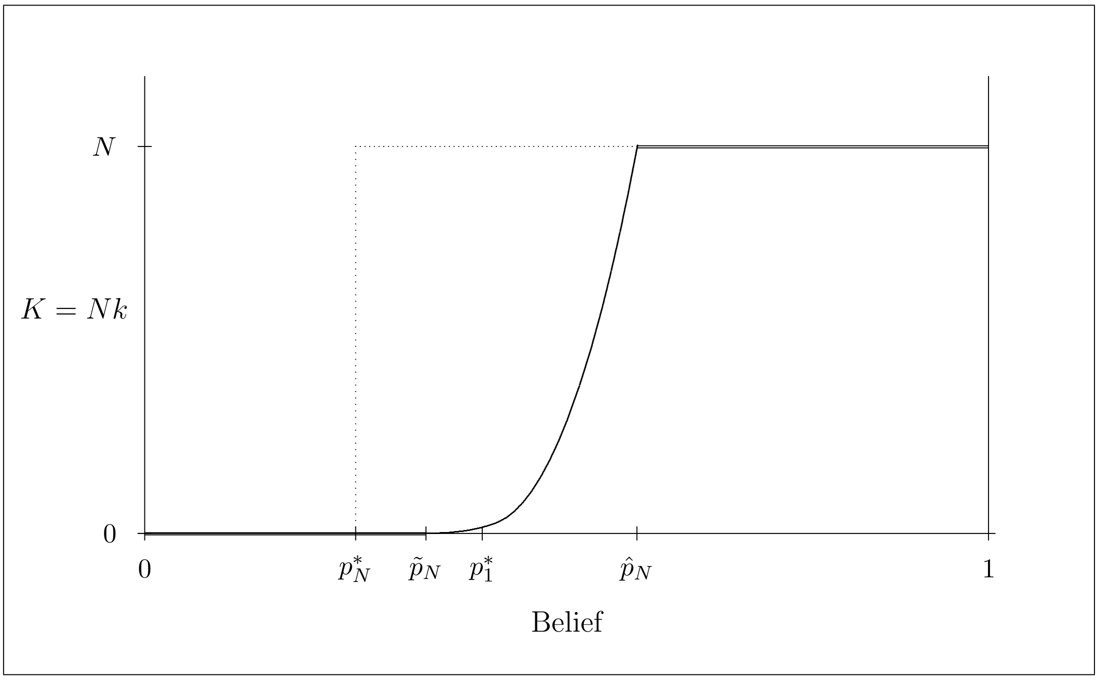


Figure 1: Intensity of experimentation in the symmetric equilibrium

Figure 1 depicts the intensity of experimentation in the symmetric equilibrium (solid curve). The dotted step function is the efficient intensity.

The symmetric equilibrium of the Poisson model shares the main features with its counterpart in the Brownian model of Bolton and Harris (1999). First, because of the incentive to free-ride, experimentation stops for good inefficiently early (the lower threshold $\tilde{p}_N$ is above the cooperative cut-off p_N^*), and the intensity of experimentation is inefficiently low at any belief between p_N^* and $p_N^\dagger$. Second, there is the encouragement effect ($\tilde{p}_N$ is below the single-agent cut-off p_1^*). Third, both the incentive to free-ride and the encouragement effect become stronger as the number of players increases.[6] Fourth, the acquisition of information is slowed down so severely near $\tilde{p}_N$ that the players' beliefs cannot reach this threshold in finite time.

Corollary 2 *Starting from a prior belief above $\tilde{p}_N$, the players' common posterior belief never reaches this threshold in the symmetric Markov perfect equilibrium.*

[6] As N increases, each player obtains a higher payoff at all beliefs where the risky arm is used some of the time, and $\tilde{p}_N$ falls. The diagonal $\mathcal{D}_{N-1}$ rotates clockwise, tending to increase $p_N^\dagger$, but since the payoff function shifts upward, the overall effect on $p_N^\dagger$ is ambiguous.

PROOF: Close to the right of $\tilde{p}_N$, the dynamics of the belief p given no success are

$$dp = -\Delta\lambda \, \frac{N}{N-1} \, \frac{W_N(p) - s}{c(p)} \, p\,(1-p)\,dt.$$

(A success merely causes a delay before the belief decays to near $\tilde{p}_N$ again.) As W_N is of class C^2 to the right of $\tilde{p}_N$ with $W_N(\tilde{p}_N) = s$, $W_N'(\tilde{p}_N) = 0$ and $W_N''(\tilde{p}_N+) \geq 0$, we can find a positive constant C such that

$$\Delta\lambda \, \frac{N}{N-1} \, \frac{W_N(p) - s}{c(p)} \, p\,(1-p) < C\,(p - \tilde{p}_N)^2$$

in a neighborhood of $\tilde{p}_N$. Starting from an initial belief $p_0 > \tilde{p}_N$ in this neighborhood, consider the dynamics $dp = -C\,(p - \tilde{p}_N)^2\,dt$. The solution with initial value p_0,

$$p_t = \tilde{p}_N + \frac{1}{Ct + (p_0 - \tilde{p}_N)^{-1}}\,,$$

does not reach $\tilde{p}_N$ in finite time. Since the modified dynamics decrease faster than the original ones, this result carries over to the true evolution of beliefs. ■

This result strongly suggests that asymmetric equilibria where a last experimenter keeps the rate of information acquisition bounded away from zero before all experimentation ceases ought to be more efficient than the symmetric one. The next section, in which we construct a variety of asymmetric MPE, confirms this conjecture.

6 Asymmetric Equilibria

We first address the possibility of finding equilibria where all players use cut-off strategies; by Proposition 3, these would necessarily be asymmetric. After deriving a lower bound on the equilibrium payoff of a last experimenter, we then construct asymmetric Markov perfect equilibria with symmetric continuation values after any success on a risky arm. We show that this can be done in a way that Pareto improves on the symmetric MPE.

6.1 Non-Existence of Equilibria in Cut-Off Strategies

In Keller, Rady and Cripps (2005), the non-existence of equilibria where all players use cut-off strategies when $\lambda_0 = 0$ emerges from the explicit construction of all Markov perfect equilibria in the two-player case. Here, we provide a direct proof.

Proposition 4 (No MPE in cut-off strategies) *In any Markov perfect equilibrium, at least one player uses a strategy that is not of the cut-off type.*

PROOF: Suppose to the contrary that there is an MPE where all players use a cut-off strategy. For $n = 1, \ldots, N$, let p_n denote the belief at which player n switches from using R exclusively to using S exclusively. Clearly, $p_n \leq p^m$ for all n. Without loss of generality, we can assume that $p_1 \leq p_2 \leq \ldots \leq p_{N-1} \leq p_N$. Moreover, we must have $p_1 < p^m$ since each player would have an incentive to deviate to the optimal strategy of a single player otherwise.

Suppose that $p_1 = p_2$. Immediately to the right of this cut-off, both u_1 and u_2 must then lie below $\mathcal{D}_1$, so players 1 and 2 playing R are not best responses. This proves that $p_1 < p_2$.

Now, u_2 must lie below $\mathcal{D}_1$ immediately to the left of p_2 (as player 2 finds it optimal to free-ride on one opponent who plays R) and above $\mathcal{D}_1$ immediately to the right of p_2 (as player 2 finds it optimal to join in with at least one opponent who plays R), so u_2 crosses $\mathcal{D}_1$ at p_2. (In fact, one can iterate this argument to establish that all cut-offs are different, and that u_n crosses $\mathcal{D}_{n-1}$ at p_n.)

Since a player's payoff function is weakly increasing in the intensity of experimentation provided by the other players, we have $u_1 \leq u_2$, and so u_1 is either below or exactly on $\mathcal{D}_1$ at p_2. In the first case, there is an interval $]p_2, p_2 + \epsilon[$ where player 1 (who is assumed to play R) is not responding optimally to the other players' combined intensity of experimentation $K_{\neg 1} = 1$. In the second case, $u_1 = u_2$ on $[p_2, 1]$ and $u_1'(p_2-) \geq u_2'(p_2-)$, hence $b(p_2, u_1) \leq b(p_2, u_2)$. But then, $u_2(p_2) = s + b(p_2, u_2) > s + b(p_2, u_1) - c(p_2) = u_1(p_2)$, a contradiction. ∎

This result forces us to construct equilibria in strategies that are more complex than cut-off strategies. The following subsection prepares the ground for this construction.

6.2 A Lower Bound on the Payoff of the Last Experimenter

We say that a Markov perfect equilibrium has a *last experimenter* if, with $\bar{p} = \inf\{p : K(p) > 0\}$, there is a player n and an $\epsilon > 0$ such that $k_n = 1$ and $K_{\neg n} = 0$ on $]\bar{p}, \bar{p} + \epsilon]$. Any simple MPE (that is, an equilibrium where all players use simple strategies as defined in Section 2) has a last experimenter. In fact, by the continuity of the players' payoff functions, all points $(p, u_n(p))$ lie below the diagonal $\mathcal{D}_1$ in (p, u)-space as p approaches $\bar{p}$ from the right, so exactly one player must be playing risky on $]\bar{p}, \bar{p} + \epsilon]$ for some $\epsilon > 0$ while all other players play safe.

The following proposition derives a lower bound on the post-success equilibrium payoff of a last experimenter.

Proposition 5 (Last experimenter) *In a Markov perfect equilibrium where the last experimenter irrevocably switches to the safe arm at belief $\bar{p}$, his payoff at the belief $j(\bar{p})$ is at least as high as that of any of his opponents.*

PROOF: Optimal behavior of the last experimenter (player 1, say) requires $c(\bar{p}) = b(\bar{p}, u_1) = \lambda(\bar{p})[u_1(j(\bar{p})) - s]/r$ as the left derivative $u_1'(\bar{p}) = 0$. If there were another player (player 2, say) with $u_2(j(\bar{p})) > u_1(j(\bar{p}))$, we would have $b(\bar{p}, u_2) = \lambda(\bar{p})[u_2(j(\bar{p})) - s]/r > c(\bar{p})$. So player 2 would act suboptimally on $[\bar{p} - \epsilon, \bar{p}]$ for some $\epsilon > 0$. ■

When $\lambda_0 = 0$, Proposition 5 still holds but does not impose any restriction because all players' values jump to the same level, $\lambda_1 h$, when a lump-sum arrives. This allows Keller, Rady and Cripps (2005) to construct simple equilibria in which the last experimenter's payoff is below his opponents' at all beliefs where the intensity of experimentation is non-zero. In particular, they provide an algorithm for the construction of the most inequitable (and least efficient) equilibrium for any number of players, and show that even this 'worst' asymmetric equilibrium dominates the symmetric one in terms of the players' average payoffs.[7]

The algorithm in Keller, Rady and Cripps (2005) exploits the absence of the encouragement effect when $\lambda_0 = 0$ and uses a backward induction approach anchored at the single-agent cut-off. In view of Propositions 2 and 5, we cannot use this approach here. The following section offers an alternative.

6.3 Constructing Asymmetric Equilibria

Our construction of asymmetric Markov perfect equilibria rests on two ideas. The first is to give the players a common continuation value after any success on a risky

[7] While Proposition 5 implies that this most inequitable equilibrium has no counterpart in the present setting, Section 7.2 below will make it clear that by rewarding the last experimenter and letting him enjoy an ever longer free-ride before the last leg, one can obtain Markov perfect equilibria with a degree of payoff asymmetry approaching that in the most inequitable MPE of Keller, Rady and Cripps (2005).

arm; the second is to let them alternate between the roles of experimenter and free-rider before all experimentation stops. Assigning symmetric continuation values after successes allows us to construct the players' average payoff function before assigning individual strategies. Letting players take turns playing risky allows us to achieve an overall intensity of experimentation higher than in the symmetric equilibrium, yielding higher equilibrium payoffs.

In fact, for points (p, u) below the diagonal $\mathcal{D}_{1-1/N}$, the common action that keeps all players indifferent between R and S and gives them u as the common continuation value, $k = (u - s)/[(N - 1)c(p)]$, implies an intensity of experimentation $K = Nk < 1$. In contrast, the equilibria we construct will have $K = 1$ over some range of beliefs where the graph of the average payoff function lies below $\mathcal{D}_{1-1/N}$. We achieve this by partitioning the range in question into a finite number of intervals on each of which exactly one player plays risky.

If the last of these 'lone experimenters' stops using the risky arm at the belief $\bar{p}$, his value function u satisfies $\lambda(\bar{p})[u(j(\bar{p})) - s]/r = c(\bar{p})$; cf. the proof of Proposition 5. When all players have a common continuation value after a success on a risky arm, this equation also holds for the players' average payoff function $\bar{u}$, and so $\lambda(\bar{p})[\bar{u}(j(\bar{p})) - s]/r = c(\bar{p})$. Varying $\bar{p}$, we can trace out the locus $\bar{\mathcal{D}}$ of all possible post-jump points $(j(\bar{p}), \bar{u}(j(\bar{p})))$ in the (p, u)-plane that satisfy this condition:

$$\bar{\mathcal{D}} = \{(p, u) \in [0, 1] \times I\!R_+ : \ \lambda(j^{-1}(p))[u - s]/r = c(j^{-1}(p))\}.$$

Using the fact that $\lambda(j^{-1}(p)) = \lambda_0\lambda_1/[p\lambda_0 + (1 - p)\lambda_1]$, it is straightforward to show that $\bar{\mathcal{D}}$ is a downward sloping straight line through the points $(0, s + r[s - \lambda_0 h]/\lambda_0)$ and $(j(p^m), s)$.

To ensure both a common continuation value after any success and an increase in the intensity of experimentation relative to the symmetric MPE, we start our construction of the average equilibrium payoff function at some point $(p^\sharp, u)$ on the lower envelope $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$ of the diagonals $\bar{\mathcal{D}}$ and $\mathcal{D}_{1-1/N}$. This lower envelope coincides with $\mathcal{D}_{1-1/N}$ if and only if $\frac{r}{\lambda_0} \geq 1 - \frac{1}{N}$, so $\bar{\mathcal{D}}$ is relevant for sufficiently high λ_0 only, that is, for sufficiently small jumps of beliefs after successes. To the right of $p^\sharp$, we proceed as in the construction of the symmetric MPE. To the left of $p^\sharp$, we solve for the average payoff function when one player out of N is playing risky. Varying $p^\sharp$, we then ensure that the average payoff hits the level s at a belief $p^\flat$ where the last experimenter is indeed indifferent between playing risky and playing safe. If the point $(p^\sharp, u)$ thus determined lies below $\bar{\mathcal{D}}$ (and hence on $\mathcal{D}_{1-1/N}$), we have $j(p^\flat) > p^\sharp$; if this point lies on $\bar{\mathcal{D}}$, we have $j(p^\flat) = p^\sharp$. In either case, a success at any belief to the right of $p^\flat$ makes the belief jump

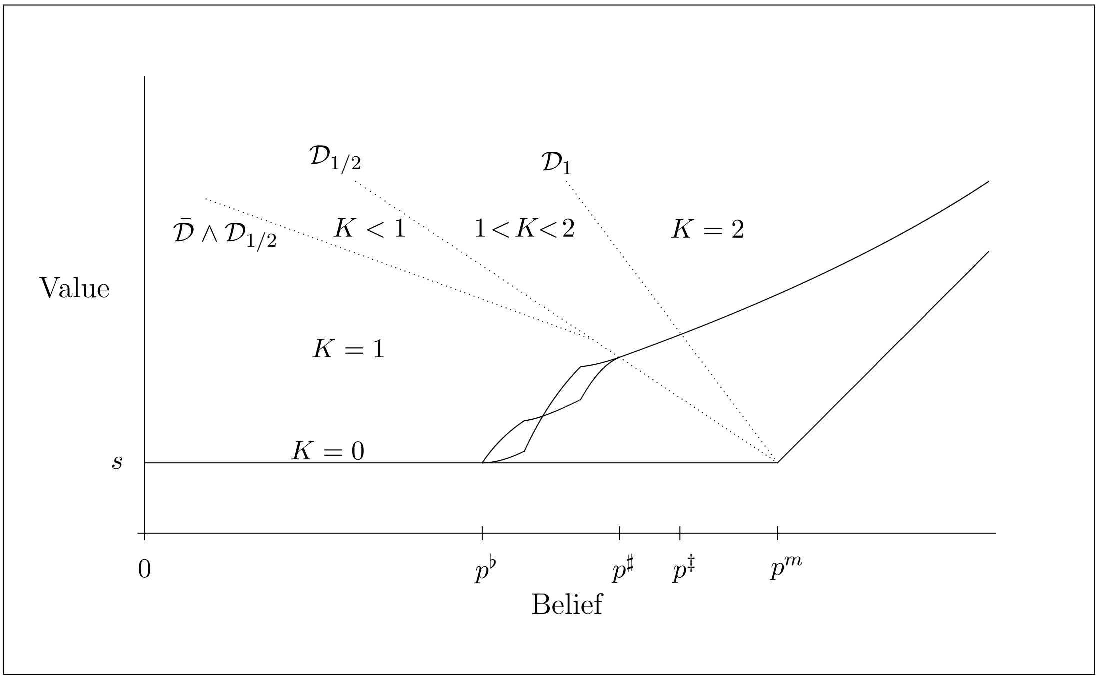


Figure 2: Intensity of experimentation in a two-player asymmetric equilibrium, and possible equilibrium payoffs

to the right of $p^\sharp$, where the equilibrium involves symmetric actions and continuation payoffs that coincide with the average. Between $p^\flat$ and $p^\sharp$, moreover, the graph of the average payoff function lies below $\mathcal{D}_{1-1/N}$, and so an intensity of experimentation equal to 1 is indeed more than would be compatible with symmetric behavior.

For $N = 2$, Figure 2 illustrates the payoff functions that can arise in the equilibria we construct and gives the corresponding intensity of experimentation in various regions of the (p, u)-plane. The faint straight lines ending in (p^m, s) are the diagonals $\mathcal{D}_1$ and $\mathcal{D}_{1/2}$, the faint straight line ending on $\mathcal{D}_{1/2}$ is the part of $\bar{\mathcal{D}} \wedge \mathcal{D}_{1/2}$ that lies below $\mathcal{D}_{1/2}$, the solid kinked line is the myopic payoff. The solid curves are the graphs of the players' payoff functions. The equilibrium intensity of experimentation varies along the graph of the average equilibrium payoff function. The intensity is 2 when the graph is above $\mathcal{D}_1$, between 1 and 2 when the graph lies between $\mathcal{D}_{1/2}$ and $\mathcal{D}_1$, etc. The intensity of experimentation is continuous in beliefs at $p^\sharp$ if the graph crosses $\bar{\mathcal{D}} \wedge \mathcal{D}_{1/2}$ on $\mathcal{D}_{1/2}$, as in the figure. If the graph crosses $\bar{\mathcal{D}} \wedge \mathcal{D}_{1/2}$ below $\mathcal{D}_{1/2}$, the intensity jumps at the belief $p^\sharp$.

For arbitrary N, we have the following result.

Proposition 6 (Asymmetric MPE) *The N-player experimentation game admits Markov perfect equilibria with three thresholds, $p_N^\flat$, $p_N^\sharp$ and $p_N^\ddagger$, where $p_N^* < p_N^\flat < p_N^\sharp < p_N^\ddagger < p^m$ and $j(p_N^\flat) \geq p_N^\sharp$, such that: on $[p_N^\sharp, 1]$, the players have a common payoff function; on $[p_N^\ddagger, 1]$, all players play R; on $]p_N^\sharp, p_N^\ddagger[$, the players allocate a common interior fraction of the unit resource to R, and this fraction increases in the belief; on $]p_N^\flat, p_N^\sharp]$, the intensity of experimentation equals 1 with players taking turns playing R on consecutive subintervals; on $[0, p_N^\flat]$, all players play S. The intensity of experimentation is continuous in beliefs on $]p_N^\flat, 1]$ with the possible exception of a jump at $p_N^\sharp$. The average payoff function is increasing on $[p_N^\flat, 1]$ and once continuously differentiable on the unit interval except for the beliefs $p_N^\flat$ and, if the intensity of experimentation has a jump there, $p_N^\sharp$. On $]p_N^\flat, 1[$, the average payoff is higher than in the symmetric MPE.*

PROOF: We just sketch the construction of the equilibrium here; details can be found in the Appendix. First, we construct the players' average payoff function $\bar{u}$ in the purported equilibria, using an approach similar to the proof of Proposition 3. This function is increasing on $[p_N^\flat, 1]$. Its graph crosses $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$ at $p_N^\sharp$ and $\mathcal{D}_{N-1}$ at $p_N^\ddagger$. It has a kink at $p_N^\sharp$ with $\bar{u}'(p_N^\sharp-) > \bar{u}'(p_N^\sharp+)$ if and only if the intersection with $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$ is below $\mathcal{D}_{1-1/N}$. It satisfies $\bar{u}(p) = s + b(p, \bar{u}) - c(p)/N$ between $p_N^\flat$ and $p_N^\sharp$, solves the indifference ODDE (6) between $p_N^\sharp$ and $p_N^\ddagger$, and is of the form (3) above $p_N^\ddagger$. The average jump benefit $\lambda(p_N^\flat)[\bar{u}(j(p_N^\flat)) - s]/r$ exactly equals the opportunity cost $c(p_N^\flat)$. As all players' payoff functions will have a zero left-hand derivative at $p_N^\flat$ and a common value of $\bar{u}(j(p_N^\flat))$ at $j(p_N^\flat)$, each player will therefore be indifferent between R and S at $p_N^\flat$.

Second, we construct the players' payoff functions and strategies. To this end, we split the interval $]p_N^\flat, p_N^\sharp]$, in finitely many subintervals $]p_{\ell,i}, p_{r,i}]$ and in turn partition each of them in N intervals $I_{1,i}, \ldots, I_{N,i}$. We let player n use R on all intervals $I_{n,i}$, and S on $]p_N^\flat, p_N^\sharp] \setminus \bigcup_i I_{n,i}$. Using intermediate-value arguments, we can choose the intervals $I_{1,i}, \ldots, I_{N,i}$ such that each player's payoff function coincides with $\bar{u}$ at the boundaries of each subinterval $]p_{\ell,i}, p_{r,i}]$. By increasing the number and reducing the size of these subintervals, moreover, we can ensure that the vertical distance $|u_n - \bar{u}|$ remains below some given real number $\delta > 0$ for all n.

Third, we verify that for sufficiently small δ, that is, for sufficiently frequent alternation between the roles of free-rider and experimenter on $]p_N^\flat, p_N^\sharp]$, the strategies we have constructed are mutually best responses.

As to the comparison of the average payoff function $\bar{u}$ with that of the symmetric equilibrium, W_N, suppose that $\bar{u} - W_N$ assumes a negative global minimum at the belief p in the open unit interval. Note that $\bar{u}$ must be differentiable there since a kink with $\bar{u}'(p^\sharp_N-) > \bar{u}'(p^\sharp_N+)$ is incompatible with even a local minimum of $\bar{u} - W_N$ at $p^\sharp_N$. At the belief p, therefore $W'_N(p) = \bar{u}'(p)$ and $\bar{u}(j(p)) - W_N(j(p)) \geq \bar{u}(p) - W_N(p)$, so $b(p, \bar{u}) \geq b(p, W_N)$. We cannot have both $W_N(p)$ and $\bar{u}(p)$ above $\mathcal{D}_1$ since in this region both W_N and $\bar{u}$ are of the form (3) and so the difference $\bar{u} - W_N$ is increasing there. Further, if $W_N(p)$ is above $\mathcal{D}_1$ and $\bar{u}(p)$ is on or below that diagonal, then $b(p, W_N) > c(p) \geq b(p, \bar{u})$ in contradiction to what we derived before (to the left of $p^\sharp_N$, $\bar{u}(p) = s + b(p, \bar{u}) - c(p)/N < s + (1 - 1/N)c(p)$ and hence $b(p, \bar{u}) < c(p)$ there). Consequently, we must have both $W_N(p)$ and $\bar{u}(p)$ on or below $\mathcal{D}_1$, which translates into $b(p, W_N) = c(p) \geq b(p, \bar{u})$ and hence, by what we saw above, $b(p, W_N) = b(p, \bar{u})$. This in turn yields $\bar{u}(j(p)) - W_N(j(p)) = \bar{u}(p) - W_N(p)$, so the difference $\bar{u} - W_N$ is also at its minimum at the belief $j(p)$. Iterating the argument until we get to the right of p^m (and hence to the right of $\mathcal{D}_1$), we again obtain a contradiction. This establishes $\bar{u} \geq W_N$. Now, if we had $\bar{u}(p) = W_N(p)$ at some $p \in]p^\flat_N, 1[$, this would again imply $b(p, \bar{u}) = b(p, W_N)$ at a global minimum of $\bar{u} - W_N$ and, by the iterative argument just given, lead to another contradiction. This proves that $\bar{u} > W_N$ on $]p^\flat_N, 1[$. In particular, $p^\flat_N$ lies to the left of the belief $\tilde{p}_N$ at which all experimentation stops in the symmetric MPE. ■

The gain in average payoffs relative to the symmetric equilibrium stems from the fact that, owing to the alternation between the roles of single experimenter and free-rider, the intensity of experimentation is bounded away from zero immediately above the belief where all experimentation stops. In the symmetric equilibrium, a player who deviates to the safe action slows down the gradual slide of beliefs towards more pessimism; as the opponents' strategies are increasing functions of the level of optimism, the deviation causes them to experiment more than they would on the equilibrium path. When players use beliefs to coordinate their alternation between experimentation and free-riding, by contrast, a deviation from the risky to the safe action freezes the belief in its current state and delays the time at which another player takes over the burden of experimentation. Deviations are thus more attractive under symmetric strategies than under alternation. This explains why the equilibrium intensity under the latter can be higher.

For beliefs above $p^\sharp_N$, the players' common payoff function permits an explicit representation of the form given in Corollary 1. For beliefs between $p^\flat_N$ and $p^\sharp_N$, we have the following result.

Corollary 3 *For $p \in]p^\flat_N, p^\sharp_N]$, let ι be the smallest integer such that $j^{\iota+1}(p) \geq p^\ddagger_N$, i.e. $\iota + 1$ consecutive successes would result in all the players playing R exclusively. Then,* *with $k_n = 1$ for an experimenter and $k_n = 0$ for a free-rider, the payoff functions are*

$$
\begin{aligned}
u_n(p) \;&=\; \lambda(p)h + (\iota + k_n - 1)\left[\tfrac{r}{r+\lambda_1}(\lambda_1 h - s)p - \tfrac{r}{r+\lambda_0}(s - \lambda_0 h)(1-p)\right] \\
&+ C^{(0)}\left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu_N}}{\lambda_0 - \mu_N\Delta\lambda}\right)^{\iota}\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu_N}}{r + \lambda_0 - \mu_N\Delta\lambda}(1-p)\,\Omega(p)^{\mu_N} \\
&+ \frac{1}{r}\left\{\sum_{\eta=0}^{\iota-1}\frac{C^{(\iota-\eta)}}{\eta!}\left(-\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\lambda_0/\Delta\lambda}}{\Delta\lambda}\right)^{\eta}\lambda_0\,(\lambda_0/\lambda_1)^{\lambda_0/\Delta\lambda}\right. \\
&\qquad\left.\times\left(\sum_{\gamma=0}^{\eta}\left(\frac{\Delta\lambda}{r}\right)^{\eta-\gamma}\frac{\eta!}{\gamma!}\left(\ln\left[(\lambda_0/\lambda_1)^{\gamma}\,\Omega(p)\right]\right)^{\gamma}\right)\right\}(1-p)\,\Omega(p)^{\lambda_0/\Delta\lambda} \\
&+ C_n^{(\iota+1)}(1-p)\,\Omega(p)^{(r+\lambda_0)/\Delta\lambda}\,,
\end{aligned}
$$

for appropriately chosen constants of integration $C_n^{(\iota+1)}$; the constants $C^{(i)}$, $i = 0, \ldots, \iota$, are from the common payoff function for beliefs above $p_N^\sharp$.

PROOF: See the Appendix. ■

6.4 Pareto Improvements over the Symmetric Equilibrium

With sufficiently frequent turns between the roles of experimenter and free-rider, the players' payoff functions in the equilibria of Proposition 6 become arbitrarily close to the average payoff function. This leads to a Pareto improvement over the symmetric equilibrium.

Proposition 7 (Pareto improvement over the symmetric MPE) *The N-player experimentation game admits Markov perfect equilibria as in Proposition 6 in which each player's payoff exceeds the symmetric equilibrium payoff on $]p_N^\flat, 1[$.*

PROOF: Let $\delta = \frac{1}{2}\max_{\tilde{p}_N \le p \le p_N^\sharp}[\bar{u}(p) - W_N(p)]$, where $\bar{u}$ is the average payoff function associated with the equilibria of Proposition 6, and W_N is the players' common payoff function in the symmetric equilibrium. Choose the subintervals $]p_{\ell,i}, p_{r,i}]$ such that $|u_n - \bar{u}|$ is bounded above by δ for all n. Then $u_n > s = W_N$ on $]p_N^\flat, \tilde{p}_N]$, $u_n \ge \bar{u} - \delta > W_N$ on $]\tilde{p}_N, p_N^\sharp]$, and $u_n = \bar{u} > W_N$ on $]p_N^\sharp, 1]$. ■

It deserves to be stressed that it is the encouragement effect which permits Pareto improvements over the symmetric equilibrium. Without it, the last experimenter quits at the same belief (the single-agent cut-off) at which all players stop experimenting in the symmetric equilibrium; bearing all the costs of experimentation on his own, the last experimenter is then necessarily worse off than under symmetry immediately to the right of this cut-off.

There clearly is scope for further improvements in players' equilibrium payoffs, over and above those embodied in the equilibria of Proposition 6. Moving down from the diagonal $\mathcal{D}_{N-2+1/N}$ to $\mathcal{D}_{N-2}$, for example, the intensity of experimentation in these equilibria gradually falls from $N-1$ to $N(N-2)/(N-1)$. Using exactly the same approach as on the interval $]p_N^\flat, p_N^\sharp]$ above, we could instead let players take turns between the roles of experimenter and free-rider such that the intensity of experimentation remained constant at level $N-1$ in between these diagonals. Similar improvements are possible at lower intensities of experimentation, but since they do not add to the insights already gained, we do not pursue them here.

7 Simple Equilibria

Keller, Rady and Cripps (2005) show how to construct a variety of simple equilibria for the exponential bandit model ($\lambda_0 = 0$). Our next aim is to investigate simple equilibria in the Poisson model as λ_0 approaches zero. The similarities and differences between the two frameworks already become apparent in the two-player case, so we will restrict our attention to this case in what follows.

We start with the observation that for λ_0 sufficiently close to zero, any success on a risky arm makes the players so optimistic that playing risky is the dominant action for all of them. More precisely, as any two-player MPE must have an average payoff function in between the single-agent optimum V_1^* and the cooperative solution V_2^*, the infimum of the set $\{p : K(p) > 0\}$ must be at least p_2^*. If $j(p_2^*) \geq p^m$, a success on any risky arm will make players optimistic enough for all of them to revert to exclusive use of the risky arm, and the players' post-jump equilibrium payoffs as well as their average will be of the form $V_2(j(p))$ with V_2 as given in (3).

A necessary and sufficient condition for $j(p_2^*) \geq p^m$ is that $\mu_2/(\mu_2 + 1) \geq \lambda_0/\lambda_1$ or, equivalently, that $\mu_2 \geq \lambda_0/\Delta\lambda$. Using (2) with $N = 2$, this holds if and only if

$$\lambda_0 \left(\frac{\lambda_0}{\lambda_1}\right)^{\frac{\lambda_0}{\Delta\lambda}} \leq \frac{r}{2}\,. \tag{7}$$

Clearly, since $\lambda_0/\lambda_1 < 1$ and $\lambda_0/\Delta\lambda > 0$, a sufficient condition for (7) is that $\lambda_0 \leq r/2$.

When (7) holds, the same approach as in the previous section allows us to construct simple equilibria.

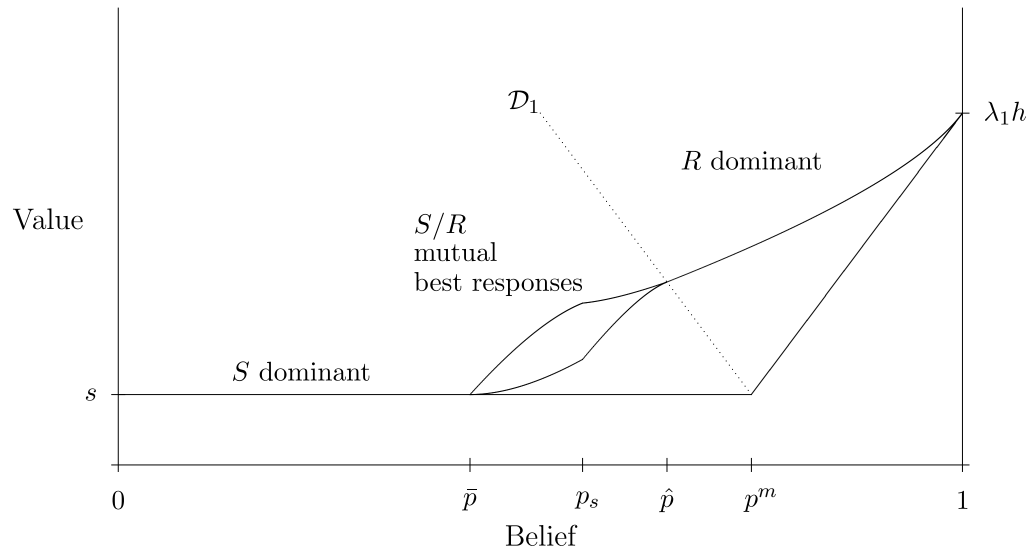


Figure 3: Best responses for $N = 2$, and possible payoffs in a simple equilibrium

7.1 Simple MPE with Common Values after a Success

Figure 3 shows the best response correspondence for $N = 2$ and illustrates the simplest possible configuration of payoff functions that can arise in the type of equilibrium we construct.

Proposition 8 (Simple MPE for $N = 2$) *Under condition* (7)*, the two-player experimentation game admits simple Markov perfect equilibria with the following features. There are two thresholds, $\bar{p}$ and $\hat{p}$, with $p_2^* < \bar{p} < \hat{p} < p^m$, such that: on $]\hat{p}, 1]$, both players play R and their payoff functions coincide; on $]\bar{p}, \hat{p}]$, the intensity of experimentation equals 1, and there is at least one belief in the interior of this interval where both players change action; on $[0, \bar{p}]$, they both play S.*

Proof: We just sketch the proof here; details can be found in the Appendix. First, we construct the two players' average payoff function $\bar{u}$ in the purported equilibria, proceeding along the same lines as in the proof of Proposition 6. This function is increasing on $[\bar{p}, 1]$. Varying the belief $\hat{p}$ at which the function crosses the diagonal $\mathcal{D}_1$, and exploiting the large-jumps condition (7), we ensure that the jump benefit $\lambda(\bar{p})[\bar{u}(j(\bar{p})) - s]/r$ exactly equals the opportunity cost $c(\bar{p})$. As both players' payoff functions will have a zero left-hand derivative at $\bar{p}$, they will thus both be indifferent between R and S at this belief.

Second, we construct the players' payoff functions and strategies. To this end, we split the interval $]\bar{p}, \hat{p}]$ in finitely many subintervals $]p_{\ell,i}, p_{r,i}]$ and select a switchpoint $p_{s,i}$ in the interior of each. On $]p_{\ell,i}, p_{s,i}]$, we let player 1 play R and player 2 play S; on $]p_{s,i}, p_{r,i}]$, we let player 1 play S and player 2 play R. An intermediate-value argument shows that the switchpoints $p_{s,i}$ can be chosen such that both players' payoff functions coincide with $\bar{u}$ at the boundaries of each subinterval. By construction, u_1 satisfies smooth pasting at $\bar{p}$. We show that $u_1 \leq \bar{u} \leq u_2$. We also establish that u_1 is increasing on $[\bar{p}, 1]$, and u_2 at least on all intervals $[p_{\ell,i}, p_{s,i}]$ and $[\hat{p}, 1]$. We cannot rule out the possibility that u_2 is decreasing on some interval $[p_{s,i}, p_{s,i} + \epsilon]$, where player 2 is the experimenter.[8]

Third, we establish that the left-hand derivative of u_2 at $\hat{p}$ satisfies $u_2'(\hat{p}) > -\Delta\lambda h$. This allows us to choose the above subintervals so that (despite a possible lack of monotonicity) u_2 stays below $\mathcal{D}_1$ to the left of $\hat{p}$; by increasing the number and reducing the size of the subintervals, moreover, we can ensure that the vertical distance $u_2 - u_1$ never exceeds some given real number $\delta > 0$.

Fourth, we verify that the strategies so constructed are mutual best responses. In view of the monotonicity of u_1, this is very easy for player 1. If u_2 is not monotonic, we obtain the best-response property for player 2 by taking δ small enough, that is, by imposing sufficiently frequent alternation between the roles of experimenter and free-rider. ■

Figure 3 illustrates the case where one switchpoint p_s in $]\bar{p}, \hat{p}[$ suffices to construct an equilibrium as just outlined with a monotonic payoff function u_2 (the higher of the two payoff functions). This case arises for example with the parameter values $r = 1, s = 1.5, h = 2, \lambda_0 = 0.5, \lambda_1 = 1.5$. With these values, $\bar{p} < \tilde{p}_2$, the belief where all experimentation stops in the symmetric equilibrium, and the average payoff function is greater than the common payoff function in the symmetric equilibrium; this improvement stems again from the fact that the intensity of experimentation remains constant at the level 1 just below $\mathcal{D}_{1/2}$, whereas the symmetric equilibrium intensity falls below 1 as soon as $\mathcal{D}_{1/2}$ is crossed.[9]

[8] For more on the non-monotonicity of u_2, see the remarks about 'anticipation' in the next subsection on numerical solutions.

[9] However, it is not the case that each player is individually better off than in the symmetric MPE – in fact, the last experimenter (the player with the lower of the two payoff functions) is worse off in a neighborhood of the switchpoint. Below, we will present simple equilibria for the above parameter

Using condition (7), we can give the following explicit representations for the two players' payoff functions in the equilibria of Proposition 8.

Corollary 4 *On $]\hat{p}, 1]$, where both players experiment, their common payoff function is*

$$u(p) = \lambda(p)h + C^{(0)}(1-p)\,\Omega(p)^{\mu_2},$$

the constant $C^{(0)}$ being given by

$$C^{(0)} = 2\,c(\hat{p})\,(1-\hat{p})^{-1}\,\Omega(\hat{p})^{-\mu_2}.$$

On $]\bar{p}, \hat{p}]$, where one player experiments and the other free-rides, and with $k_n = 1$ for an experimenter and $k_n = 0$ for a free-rider, the payoff functions are

$$\begin{aligned} u_n(p) &= \lambda(p)h + (k_n - 1)\left[\tfrac{r}{r+\lambda_1}(\lambda_1 h - s)p - \tfrac{r}{r+\lambda_0}(s - \lambda_0 h)(1-p)\right] \\ &\quad + C^{(0)}\left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu_2}}{r + \lambda_0 - \mu_2\Delta\lambda}\right)(1-p)\,\Omega(p)^{\mu_2} + C_n^{(1)}(1-p)\,\Omega(p)^{(r+\lambda_0)/\Delta\lambda}, \end{aligned}$$

with appropriately chosen constants of integration $C_n^{(1)}$.

PROOF: See the Appendix. ■

For $\lambda_0 = 0$, the results of Keller, Rady and Cripps (2005) imply that equilibria as in Proposition 8 generate, at any belief, the highest average payoff achievable in a simple two-player MPE with a last experimenter. As the following subsection shows, this result does not generalize to $\lambda_0 > 0$.

7.2 Rewarding the Last Experimenter

So far, we have constructed equilibria where the continuation value immediately after a success on any risky arm is the same for both players, which means in particular that both players are indifferent between R and S at the belief $\bar{p}$ where the last experimenter quits. Maintaining assumption (7), suppose now we give the last experimenter a higher payoff at $j(\bar{p})$ than the other player. This has two effects. On the one hand, we can no longer achieve the maximal intensity of 2 immediately to the right of the belief at which the graph of the average payoff function crosses $\mathcal{D}_1$, which lowers average payoffs. On the other hand, the last experimenter is now willing to continue playing R

values that are better than the symmetric one for *both* players.

somewhat to the left of $\bar{p}$, which increases average payoffs. Our final goal is to explore this trade-off numerically.

We refer to the last experimenter as player 1 and the last free-rider as player 2, and continue to let $\bar{p}$ denote the belief where all experimentation stops, and p_s the switchpoint where both players change actions; however, as their payoff functions cross $\mathcal{D}_1$ at different points, we let $\hat{p}_1$ and $\hat{p}_2$ denote the corresponding beliefs.

Construction of equilibria

The strategies for the players are: on $]\hat{p}_2, 1]$ both players play R; on $]\hat{p}_1, \hat{p}_2]$ player 1 plays R and player 2 plays S; on $]p_s, \hat{p}_1]$ player 1 plays S and player 2 plays R; on $]\bar{p}, p_s]$ player 1 plays R and player 2 plays S; on $[0, \bar{p}]$ both players play S. We need to determine $\bar{p} < p_s < \hat{p}_1 < \hat{p}_2$, and to build continuous functions u_1 and u_2 that (a) connect the points $(0, s)$ and $(1, \lambda_1 h)$ in the (p, u)-plane, that (b) satisfy the appropriate ODDEs, and that (c) have the following properties: u_1 is above $\mathcal{D}_1$ on $]\hat{p}_1, 1]$, below $\mathcal{D}_1$ but above s on $]\bar{p}, \hat{p}_1]$, at s on $[0, \bar{p}]$, and is smooth at $\bar{p}$; u_2 is above $\mathcal{D}_1$ on $]\hat{p}_2, 1]$, below $\mathcal{D}_1$ but above s on $]\bar{p}, \hat{p}_2]$, and at s on $[0, \bar{p}]$.

Relative to the upper threshold $\hat{p}$ and the average payoff function $\bar{u}$ from Proposition 8, choose a point $(\hat{p}_2, \check{u})$ in $[0, 1] \times [s, \lambda_1 h]$ with $\hat{p}_2$ to the right of $\hat{p}$ and $\check{u}$ above $\bar{u}(\hat{p}_2)$. First, we construct player 1's payoff function piecewise.

On $]\hat{p}_2, 1]$ both players play R, so u_1 is of the form given in equation (3) with $N = 2$, the constant of integration being chosen so that $u_1(\hat{p}_2) = \check{u}$. The belief $\bar{p}$ where player 1 quits can now be determined from equation (1) with $N = 1$, using value matching ($u_1(\bar{p}) = s$) and smooth pasting ($u_1'(\bar{p}) = 0$), *and* knowing the form of $u_1(j(\bar{p}))$ since $j(\bar{p}) > \hat{p}_2$. Just to the left of $\hat{p}_2$, player 1 is the only one playing R, so u_1 is of the form given in Corollary 4, the constant of integration being chosen to ensure continuity at $\hat{p}_2$; $\hat{p}_1$ is the belief to the left of $\hat{p}_2$ where u_1 crosses $\mathcal{D}_1$. On an interval to the left of that, player 1 is free-riding, u_1 is again of the form given in Corollary 4, and the constant of integration is chosen to ensure continuity at $\hat{p}_1$. Further, on an interval to the right of $\bar{p}$, player 1 is again the lone experimenter and now the constant of integration is chosen to ensure that $u_1(\bar{p}) = s$. Player 1's switchpoint is where the graph of u_1 coming down and to the left from $(\hat{p}_1, u_1(\hat{p}_1))$ intersects the curve going up and to the right from $(\bar{p}, s)$.

Player 2's payoff function is also constructed piecewise. On $]\hat{p}_2, 1]$, u_2 is also of the form given in equation (3) with $N = 2$, the constant of integration being chosen so that $u_2(\hat{p}_2)$ is on $\mathcal{D}_1$. Between $\bar{p}$ and $\hat{p}_2$, u_2 is of the form given in Corollary 4: player 2 is free-riding on the interval $]\hat{p}_1, \hat{p}_2]$ and we ensure continuity at $\hat{p}_2$; on an interval to the left of that, player 2 is the lone experimenter and we ensure continuity at $\hat{p}_1$; on an interval to the right of $\bar{p}$, player 2 free-rides and the constant of integration is chosen to ensure that $u_2(\bar{p}) = s$. Player 2's switchpoint is where the graph of u_2 coming down and to the left from $(\hat{p}_1, u_2(\hat{p}_1))$ intersects the curve going up and to the right from $(\bar{p}, s)$.

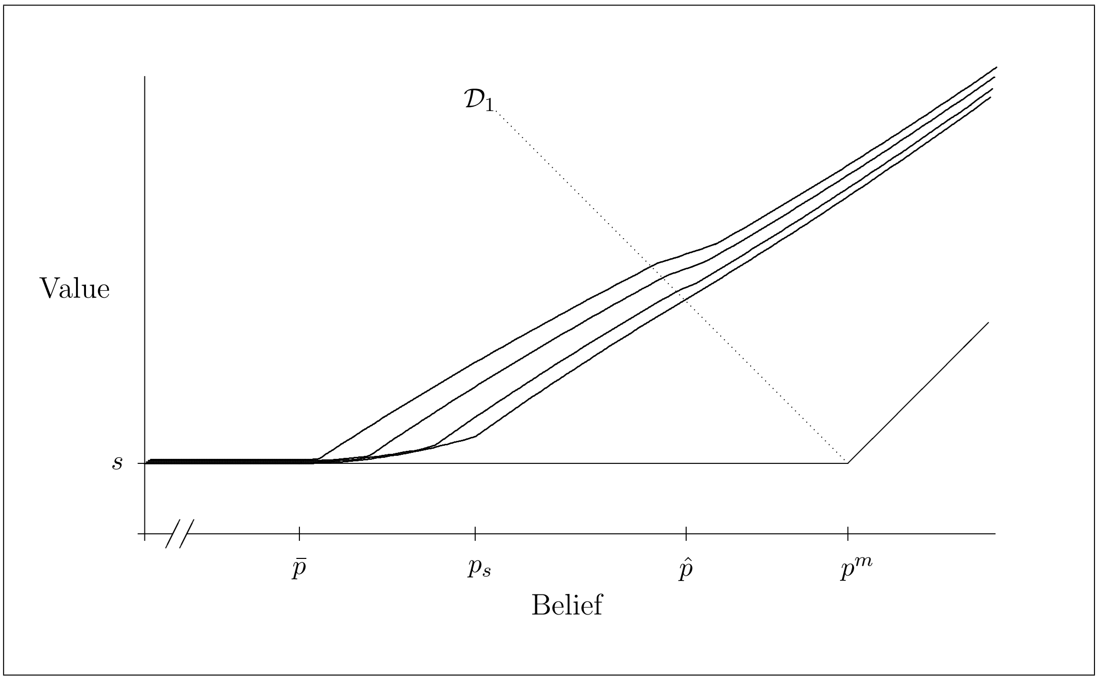
For this to be an equilibrium we need to have the players switching at the same belief – this involves adjusting $(\hat{p}_2, \check{u})$ and iterating until the switchpoints coincide.

Findings

Using the same parameter values we referred to in the discussion of Figure 3 after Proposition 8 (namely, $r = 1, s = 1.5, h = 2, \lambda_0 = 0.5, \lambda_1 = 1.5$), we numerically solved for six equilibria as well as the simple one with common payoffs above $\mathcal{D}_1$ (the 'base case'), giving the last experimenter progressively higher payoffs above $\mathcal{D}_1$. Figure 4 illustrates the players' payoffs in the base case and in three of these equilibria, the tick labels on the belief axis being for the base case (which exhibits the lowest equilibrium payoffs for the last experimenter and highest for the last free-rider.)

We find that as we improve player 1's post-jump payoff, the fall in $\bar{p}$ is about 5 times smaller than the shifts in $\hat{p}_1$ and $\hat{p}_2$, with $\hat{p}_1$ moving to the left and $\hat{p}_2$ moving to the right. (The drop in the switchpoint is more dramatic, being about 25 times that of $\bar{p}$.) The net effect is that the interval of beliefs where exactly one player is experimenting widens, and although the average payoff is higher on an interval to the right of $\bar{p}$, it dips below that of the base case very close to $\hat{p}_1$ and remains there at all higher beliefs. Player 1 is progressively better off than in the base case at all beliefs to the right of $\bar{p}$, but player 2 is progressively worse off at beliefs greater than approximately p_s, and the absolute differences between the average payoff in the base case and those in the other six equilibria become more pronounced as the asymmetry increases. Indeed, in the two most asymmetric of the equilibria that we calculated, player 2's payoff function is below the payoff function in the symmetric equilibrium in a neighborhood of $\hat{p}_1$, whereas the payoffs in the other four intermediate equilibria are Pareto improvements on the symmetric equilibrium. Moreover, in the most asymmetric of these equilibria, player 2's payoff function is *decreasing* in an interval immediately to the right of the switchpoint (although this is hard to discern visually).

Put another way, in this very asymmetric equilibrium as the players approach the switchpoint from the right, where player 2 is experimenting alone, beliefs are becoming



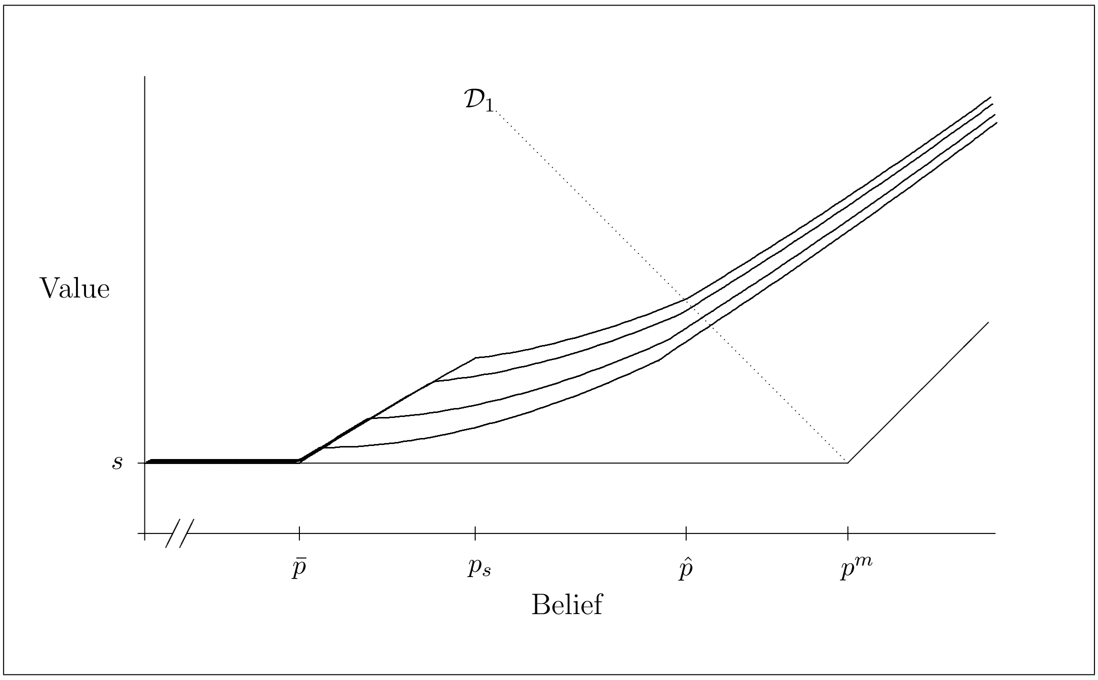


Figure 4: Equilibrium payoffs of the last experimenter (upper panel) and the last free-rider (lower panel)

more pessimistic, yet player 2's payoff is going up – if we put this down to the fact that, conditional on no impending success, player 2 will soon be able to enjoy a free-ride, then we can call this an 'anticipation effect'.[10]

8 Concluding Remarks

The asymmetric equilibria that we constructed in the Poisson framework raise the question whether similar equilibria exist in the Brownian model of Bolton and Harris (1999). The elementary constructive method that we used here is likely to apply to the Brownian case as well. Our proof of the result that there exist no equilibria in cut-off strategies should also carry over. We intend to explore this in future work.

Our model can easily be adapted to situations where an event is bad news: a 'breakdown' rather than a 'breakthrough'. For example, we can interpret s as the expected flow cost of keeping the current safe machine running. Players have access to new risky machines that break down and thereby cause lump-sum costs at exponentially distributed times; a high failure rate of λ_1 would favor the old machine, a low rate of $\lambda_0 < \lambda_1$ the new one. The aim is to minimize the expected sum of discounted costs of breakdowns. The longer the machines do not fail, the more optimistic the players become about their reliability, but whenever one does fail, the belief jumps to more pessimistic levels. Given enough pessimism, another failure will be the 'last straw'. Thus, the continuous part of the belief dynamics always keeps the state variable in the continuation region where at least some player uses the new machine, whereas the discontinuous part can cause the state variable to *jump* into the stopping region. As a consequence, the principle of smooth pasting does not apply. Despite the superficial symmetry between the 'good-news' and the 'bad-news' versions of the model, therefore, the formal analysis of the single-agent optimum, the efficient benchmark and best responses is rather different from that in the present paper. We defer such analysis to a separate paper.

By constructing simple asymmetric equilibria, our work also prepares the ground for an analysis of strategic experimentation by asymmetric players who might differ for example with respect to their innate abilities to achieve breakthroughs, the average size of lump-sum payoffs, or their outside options. This is again left to future work.

[10]This effect is also evident in the equilibria with infinitely many switches in Keller, Rady and Cripps (2005); cf. their Proposition 6.4 and Figure 3.

Appendix

Proof of Proposition 3: Let $\hat{p}_{N,N-1}$ denote the belief where the graph of V_N^* cuts $\mathcal{D}_{N-1}$, and $\hat{p}_{1,N-1}$ denote the belief where the graph of V_1^* cuts $\mathcal{D}_{N-1}$. By continuity, there is an open interval $I \supset [\hat{p}_{N,N-1}, \hat{p}_{1,N-1}]$ such that for all $\hat{p} \in I$, the unique solution to (1) that crosses $\mathcal{D}_{N-1}$ at the belief $\hat{p}$ has positive slope there.

Fix a belief $\hat{p} \in I$ and let $(\hat{p}, \hat{u})$ be the corresponding point on the diagonal $\mathcal{D}_{N-1}$. On $[\hat{p}, 1]$, we define $u^{(0)}$ as the unique solution to (1) that assumes the value $\hat{u}$ at belief $\hat{p}$. Now consider the ordinary differential equation

$$\Delta\lambda\, p(1-p)u'(p) + \lambda(p)u(p) = r\lambda(p)h - rs + \lambda(p)u^{(0)}(j(p)). \tag{A.1}$$

Standard results imply that this ODE has a unique solution $u^{(1)}$ on $[j^{-1}(\hat{p}), \hat{p}]$ with $u^{(1)}(\hat{p}) = u^{(0)}(\hat{p})$ and, by construction, $(u^{(1)})'(\hat{p}) = (u^{(0)})'(\hat{p})$.

Iterating this step, we construct functions $u^{(i+1)}$ defined on $[j^{-(i+1)}(\hat{p}), j^{-i}(\hat{p})]$ for $i = 1, 2, 3, \ldots$ by choosing $u^{(i+1)}$ as the unique solution of the ODE

$$\Delta\lambda\, p(1-p)u'(p) + \lambda(p)u(p) = r\lambda(p)h - rs + \lambda(p)u^{(i)}(j(p)) \tag{A.2}$$

subject to the condition $u^{(i+1)}(j^{-i}(\hat{p})) = u^{(i)}(j^{-i}(\hat{p}))$. Setting $u_{\hat{p}}(p) = u^{(i)}(p)$ whenever $j^{-(i+1)}(\hat{p}) \le p < j^{-i}(\hat{p})$, we thus obtain a function $u_{\hat{p}}$ of class C^1 on $]0,1]$ that solves (6) to the left of $\hat{p}$, and (1) to the right of $\hat{p}$. Standard results imply that $u_{\hat{p}}$ depends in a continuous fashion on $\hat{p}$. In particular, $M(\hat{p})$, the minimum of $u_{\hat{p}}$ on $[p_N^*, p^m]$, is continuous in $\hat{p}$.

For $\hat{p} \in I$ with $\hat{p} < \hat{p}_{N,N-1}$, the function $u_{\hat{p}}$ lies above V_N^* on at least $[\hat{p}, 1[$. If $u_{\hat{p}}$ and V_N^* assumed the same value at some belief $p_\ell \in [p_N^*, \hat{p}[$, then the restriction of $u_{\hat{p}} - V_N^*$ to $[p_\ell, 1]$ would have a positive global maximum at some belief $p_r \in]p_\ell, 1[$. In fact, we would have $p_r \in]p_\ell, \hat{p}[$ since $u_{\hat{p}} - V_N^*$, being the difference of two functions of the form (3), has a negative first derivative on $[\hat{p}, 1[$. As $(u_{\hat{p}})'(p_r) = (V_N^*)'(p_r)$ and $u_{\hat{p}}(j(p_r)) - V_N^*(j(p_r)) \le u_{\hat{p}}(p_r) - V_N^*(p_r)$, we would thus have $b(p_r, V_N^*) \ge b(p_r, u_{\hat{p}}) = c(p_r)$, hence $V_N^*(p_r) = s + Nb(p_r, V_N^*) - c(p_r) \ge s + (N-1)c(p_r)$, which is inconsistent with the fact that V_N^* is below $\mathcal{D}_{N-1}$ at p_r. Consequently, $u_{\hat{p}}$ lies above V_N^* on $[p_N^*, 1[$.

By continuity, $\hat{u}_N$, the function $u_{\hat{p}}$ obtained for $\hat{p} = \hat{p}_{N,N-1}$, lies weakly above V_N^* on $[p_N^*, 1]$. While the two functions are identical on $[\hat{p}_{N,N-1}, 1]$ by construction, they cannot be identical on the whole of $[p_N^*, \hat{p}_{N,N-1}[$ as V_N^* does not solve (A.1) immediately to the left of $\hat{p}_{N,N-1}$, for example. Arguing exactly as in the previous paragraph, we see that the restriction of $\hat{u}_N - V_N^*$ to $[p_N^*, 1]$ must assume its positive global maximum at p_N^*. This establishes $\hat{u}_N(p_N^*) > V_N^*(p_N^*) = s$. As $V_N^*(p) > s$ for $p > p_N^*$, we thus have $\hat{u}_N > s$ on $[p_N^*, 1]$, hence $M(\hat{p}_{N,N-1}) > s$.

For $\hat{p} \in I$ with $\hat{p} > \hat{p}_{1,N-1}$, the function $u_{\hat{p}}$ lies below V_1^* in a neighborhood of $\hat{p}$. If $u_{\hat{p}}$ and V_1^* assumed the same value at some belief $p_\ell \in [p_1^*, \hat{p}[$, then the restriction of $V_1^* - u_{\hat{p}}$ to $[p_\ell, 1]$ would have a positive global maximum at a belief $p_r \in]p_\ell, 1[$. As $(V_1^*)'(p_r) = (u_{\hat{p}})'(p_r)$ and $V_1^*(j(p_r)) - u_{\hat{p}}(j(p_r)) \le V_1^*(p_r) - u_{\hat{p}}(p_r)$, we would thus have $b(p_r, u_{\hat{p}}) \ge b(p_r, V_1^*)$. As

$s < V_1^*(p_r) = s + b(p_r, V_1^*) - c(p_r)$, this would imply $b(p_r, u_{\hat{p}}) > c(p_r)$ and $p_r > \hat{p}$. But then $u_{\hat{p}}(p_r) = s + Nb(p_r, u_{\hat{p}}) - c(p_r) > s + b(p_r, V_1^*) - c(p_r) = V_1^*(p_r)$, which is a contradiction. Consequently, $u_{\hat{p}}$ lies below V_1^* on $[p_1^*, \hat{p}]$.

By continuity, $\hat{u}_{1,N}$, the function $u_{\hat{p}}$ obtained for $\hat{p} = \hat{p}_{1,N-1}$, lies weakly below V_1^* on $[p_1^*, \hat{p}_{1,N-1}]$. While the two functions are identical at $\hat{p}_{1,N-1}$ by construction, they cannot be identical on the whole of $[p_1^*, \hat{p}_{1,N-1}[$. Arguing exactly as in the previous paragraph, we see that the restriction of $V_1^* - \hat{u}_{1,N}$ to $[p_1^*, 1]$ must assume its positive global maximum at p_1^*. In particular, $\hat{u}_{1,N}(p_1^*) < V_1^*(p_1^*) = s$, hence $M(\hat{p}_{1,N-1}) < s$.

So there exists a $p_N^\dagger \in \,]\hat{p}_{N,N-1}, \hat{p}_{1,N-1}[$ such that $M(p_N^\dagger) = s$. With $u^\dagger$ denoting the solution $u_{\hat{p}}$ corresponding to $\hat{p} = p_N^\dagger$, let $\tilde{p}_N$ be the highest belief in $[p_N^*, p^m]$ at which $u^\dagger$ assumes the value s. By construction, $\tilde{p}_N < p_N^\dagger < p^m$. Define the function W_N by $W_N(p) = s$ on $[0, \tilde{p}_N]$ and by $W_N(p) = u^\dagger(p) > s$ on $]\tilde{p}_N, 1]$. This is the common payoff function when all players use the strategy k described in the proposition. As a consequence, $W_N \leq V_N^*$ and in particular $\tilde{p}_N \geq p_N^*$.

If we had $\tilde{p}_N = p_N^*$, then $W_N(p_N^*) = s = V_N^*(p_N^*)$, $W_N(j(p_N^*)) \leq V_N^*(j(p_N^*))$ and $W_N'(p_N^*-) = 0 = (V_N^*)'(p_N^*)$, implying $b(p_N^*, V_N^*) \geq b(p_N^*, W_N)$. As $b(p_N^*, V_N^*) = c(p_N^*)/N$, $b(p_N^*, W_N) = c(p_N^*)$ and $c(p_N^*) > 0$, this is a contradiction. So we have $p_N^* < \tilde{p}_N < p^m$, hence $W_N'(\tilde{p}_N+) = (u^\dagger)'(\tilde{p}_N) = 0$ because the minimum of $u^\dagger$ on $[p_N^*, p^m]$ is achieved at an interior point. Thus, the function W_N is of class C^1.

It is straightforward to check from the explicit representation of W_N above $\mathcal{D}_{N-1}$ that this function is convex and increasing on $[p_N^\dagger, 1]$. Suppose W_N is not increasing on $[\tilde{p}_N, p_N^\dagger]$. Then it must assume both a local maximum and a local minimum in the interior of that interval, and there exist beliefs $p' < p''$ in $]\tilde{p}_N, p_N^\dagger[$ such that $W_N'(p') = W_N'(p'') = 0$, $W_N(p') \geq W_N(p'')$, and W_N is weakly decreasing on $[p', p'']$ and increasing on $[p'', 1]$. We now have $b(p', W_N) = \lambda(p')[W_N(j(p')) - W_N(p')]/r = c(p') > 0$, hence $W_N(j(p')) > W_N(p')$ and $j(p') > p''$. As a consequence, $W_N(j(p'')) > W_N(j(p'))$ and $b(p'', W_N) = \lambda(p'')[W_N(j(p'')) - W_N(p'')]/r > \lambda(p')[W_N(j(p')) - W_N(p')]/r = c(p') > c(p'')$, which is a contradiction. This establishes that W_N is increasing on $[\tilde{p}_N, 1]$, and k is increasing on $[\tilde{p}_N, p_N^\dagger]$.

We thus have $b(p, W_N) > c(p)$ on $]p_N^\dagger, 1]$, $b(p, W_N) = c(p)$ on $[\tilde{p}_N, p_N^\dagger]$, and, because of the monotonicity of W_N on $[\tilde{p}_N, 1]$, $b(p, W_N) < c(p)$ on $[0, \tilde{p}_N[$. So all players using the strategy k constitutes an equilibrium. Finally, $\tilde{p}_N < p_1^*$ by Proposition 2.

Uniqueness has already been shown in the main text. ■

PROOF OF COROLLARY 1: With $u^{(0)}(p) = \lambda_1 hp + \lambda_0 h(1-p) + C^{(0)}\,(1-p)\,\Omega(p)^\mu$ (see (3)), we seek a sequence of functions $u^{(i+1)}$ for $i = 0, 1, \ldots$, defined recursively as solutions to the ODE (A.2). Let $\alpha = \lambda_0/\Delta\lambda$, and, for $i \geq 0$, let

$$u^{(i)}(p) = d_1^{(i)} p + d_0^{(i)}(1-p) + m^{(i)}(1-p)\Omega(p)^\mu + (1-p)\Omega(p)^\alpha \sum_{\eta=0}^{i-1} l^{(i-\eta)} \left(\ln\left[(\lambda_0/\lambda_1)^{\eta-1}\,\Omega(p)\right]\right)^\eta$$

where $d_1^{(i)}, d_0^{(i)}, m^{(i)}, l^{(i-\eta)}$ are constants to be determined – we will show that the functions

$u^{(i)}$ form just such a sequence. Clearly we need

$$d_1^{(0)} = \lambda_1 h, \quad d_0^{(0)} = \lambda_0 h, \quad \text{and} \quad m^{(0)} = C^{(0)}$$

with $C^{(0)}$ being the constant that fixes payoffs above the diagonal where everyone plays R. The final (summed) term in the above equation defining $u^{(i)}$ is vacuous for $i = 0$.

First note that

$$\begin{aligned} u^{(i)}(j(p)) &= d_1^{(i)} \frac{\lambda_1}{\lambda(p)} p + d_0^{(i)} \frac{\lambda_0}{\lambda(p)} (1-p) + m^{(i)} \frac{\lambda_0}{\lambda(p)} \left(\frac{\lambda_0}{\lambda_1}\right)^{\mu} (1-p)\Omega(p)^{\mu} \\ &\quad + \frac{\lambda_0}{\lambda(p)} \left(\frac{\lambda_0}{\lambda_1}\right)^{\alpha} (1-p)\Omega(p)^{\alpha} \sum_{\eta=0}^{i-1} l^{(i-\eta)} \left(\ln\left[(\lambda_0/\lambda_1)^{\eta}\,\Omega(p)\right]\right)^{\eta}, \end{aligned}$$

so that the right-hand side of (A.2) becomes

$$G^{(i)}(p) = D_1^{(i)} p + D_0^{(i)}(1-p) + M^{(i)}(1-p)\Omega(p)^{\mu} + (1-p)\Omega(p)^{\alpha} \sum_{\eta=0}^{i-1} L^{(i-\eta)} \left(\ln\left[(\lambda_0/\lambda_1)^{\eta}\,\Omega(p)\right]\right)^{\eta}$$

where

$$D_1^{(i)} = d_1^{(i)}\lambda_1 + r(\lambda_1 h - s), \quad D_0^{(i)} = d_0^{(i)}\lambda_0 - r(s - \lambda_0 h)$$

and

$$M^{(i)} = m^{(i)}\lambda_0 \left(\lambda_0/\lambda_1\right)^{\mu}, \quad L^{(i-\eta)} = l^{(i-\eta)}\lambda_0 \left(\lambda_0/\lambda_1\right)^{\alpha}.$$

The homogeneous equation, $\Delta\lambda\, p(1-p)u'(p) + \lambda(p)u(p) = 0$, has the solution

$$u_0(p) = (1-p)\Omega(p)^{\alpha}.$$

Using the method of variation of constants, we now write $u(p) = a(p)u_0(p)$ so that

$$\Delta\lambda\, p(1-p)u'(p) + \lambda(p)u(p) = \Delta\lambda\, p(1-p)u_0(p)a'(p).$$

The ODE thus transforms into the following equation for the first derivative of the unknown function a:

$$\begin{aligned} \Delta\lambda\, a'(p) &= \frac{G^{(i)}(p)}{p(1-p)\, u_0(p)} \\ &= D_1^{(i)}\, \Omega(p)^{-\alpha}(1-p)^{-2} + D_0^{(i)}\, \Omega(p)^{-\alpha+1}(1-p)^{-2} + M^{(i)}\, \Omega(p)^{\mu-\alpha+1}(1-p)^{-2} \\ &\quad + \Omega(p)(1-p)^{-2} \sum_{\eta=0}^{i-1} L^{(i-\eta)} \left(\ln\left[(\lambda_0/\lambda_1)^{\eta}\,\Omega(p)\right]\right)^{\eta}. \end{aligned}$$

Make the substitution $\omega = \Omega(p)$ and define $A(\omega) = a(p)$, so $a'(p) = -A'(\omega)/p^2$. Then

$$-\Delta\lambda\, A'(\omega) = D_1^{(i)}\, \omega^{-\alpha-2} + D_0^{(i)}\, \omega^{-\alpha-1} + M^{(i)}\, \omega^{\mu-\alpha-1} + \omega^{-1} \sum_{\eta=0}^{i-1} L^{(i-\eta)} \left(\ln\left[(\lambda_0/\lambda_1)^{\eta}\,\omega\right]\right)^{\eta},$$

so

$$\begin{aligned} A(\omega) &= \frac{D_1^{(i)}}{\lambda_1}\,\omega^{-\alpha-1} + \frac{D_0^{(i)}}{\lambda_0}\,\omega^{-\alpha} + \frac{M^{(i)}}{\lambda_0 - \mu\Delta\lambda}\,\omega^{\mu-\alpha} \\ &\quad - \sum_{\eta=0}^{i-1} \frac{L^{(i-\eta)}}{(\eta+1)\Delta\lambda}\,(\ln\left[(\lambda_0/\lambda_1)^{\eta}\,\omega\right])^{\eta+1} + C^{(i+1)}, \end{aligned}$$

where $C^{(i+1)}$ is a constant of integration (and assuming $\mu \neq \alpha$, else we have another logarithmic term). Multiplying by $u_0(p) = (1-p)\,\omega^{\alpha}$ and substituting $\omega = \Omega(p)$ leads to

$$\begin{aligned} u^{(i+1)}(p) &= \frac{D_1^{(i)}}{\lambda_1}\,p + \frac{D_0^{(i)}}{\lambda_0}\,(1-p) + \frac{M^{(i)}}{\lambda_0 - \mu\Delta\lambda}\,(1-p)\Omega(p)^{\mu} \\ &\quad + (1-p)\Omega(p)^{\alpha} \sum_{\eta=1}^{i} \frac{-L^{(i+1-\eta)}}{\eta\,\Delta\lambda} \left(\ln\left[(\lambda_0/\lambda_1)^{\eta-1}\,\Omega(p)\right]\right)^{\eta} + (1-p)\Omega(p)^{\alpha} C^{(i+1)}. \end{aligned}$$

The above iterative step shows that

$$d_1^{(i+1)} = d_1^{(i)} + \tfrac{r}{\lambda_1}(\lambda_1 h - s), \quad d_0^{(i+1)} = d_0^{(i)} - \tfrac{r}{\lambda_0}(s - \lambda_0 h), \quad \text{and} \quad m^{(i+1)} = m^{(i)} \left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu}}{\lambda_0 - \mu\Delta\lambda}\right)$$

and so, in general,

$$d_1^{(i)} = \lambda_1 h + \tfrac{r}{\lambda_1}(\lambda_1 h - s)\,i\,, \quad d_0^{(i)} = \lambda_0 h - \tfrac{r}{\lambda_0}(s - \lambda_0 h)\,i\,, \quad \text{and} \quad m^{(i)} = C^{(0)} \left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu}}{\lambda_0 - \mu\Delta\lambda}\right)^{i}.$$

After a little algebra, we find that the constants in the summation are given by

$$l^{(i-\eta)} = \frac{C^{(i-\eta)}}{\eta!} \left(-\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\alpha}}{\Delta\lambda}\right)^{\eta} \quad \text{for } \eta = 0, \ldots, i-1.$$

The constants $C^{(i-\eta)}$ $(\eta = 0, \ldots i-1)$ are chosen to ensure continuity. In particular, writing $\hat{\jmath}^{-i}$ for $j^{-i}(p_N^{\dagger})$, $C^{(i+1)}$ is chosen such that $u^{(i+1)}(\hat{\jmath}^{-i}) = u^{(i)}(\hat{\jmath}^{-i})$ for $i \geq 0$, and satisfies

$$\begin{aligned} &C^{(i+1)}\,(1-\hat{\jmath}^{-i})\Omega(\hat{\jmath}^{-i})^{\alpha} \\ &= -\tfrac{r}{\lambda_1}(\lambda_1 h - s)\,\hat{\jmath}^{-i} + \tfrac{r}{\lambda_0}(s - \lambda_0 h)\,(1-\hat{\jmath}^{-i}) \\ &\quad + C^{(0)} \left(1 - \frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu}}{\lambda_0 - \mu\Delta\lambda}\right) \left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu}}{\lambda_0 - \mu\Delta\lambda}\right)^{i} (1-\hat{\jmath}^{-i})\Omega(\hat{\jmath}^{-i})^{\mu} \\ &\quad + \left\{ \sum_{\eta=0}^{i-1} C^{(i-\eta)} \left[\frac{1}{\eta!} \left(-\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\alpha}}{\Delta\lambda} \ln\left[(\lambda_0/\lambda_1)^{\eta-1}\,\Omega(\hat{\jmath}^{-i})\right]\right)^{\eta} \right.\right. \\ &\qquad \left.\left. - \frac{1}{(\eta+1)!} \left(-\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\alpha}}{\Delta\lambda} \ln\left[(\lambda_0/\lambda_1)^{\eta}\,\Omega(\hat{\jmath}^{-i})\right]\right)^{\eta+1} \right] \right\} (1-\hat{\jmath}^{-i})\Omega(\hat{\jmath}^{-i})^{\alpha}. \end{aligned}$$

■

Proof of Corollary 3: The proof follows the same lines as in Corollary 1. Here we consider the case where exactly 1 of the N players is playing R, and where a success results in the players playing symmetrically as in that corollary. The relevant ODE is

$$\Delta\lambda\, p(1-p)u'(p) + [r+\lambda(p)]u(p) = kr\lambda(p)h + (1-k)rs + \lambda(p)w(j(p))$$

with $k=0$ for a free-rider and $k=1$ for an experimenter, and where w is the function $u^{(\iota)}$ derived in Corollary 1.

We obtain equations for u_n of the form

$$\begin{aligned} u_n(p) \;=\;& d_1^{(\iota+1)}(k_n)\, p + d_0^{(\iota+1)}(k_n)\,(1-p) + m^{(\iota+1)}\,(1-p)\Omega(p)^{\mu_N} \\ &+ \frac{1}{r}\sum_{\eta=0}^{\iota-1} l^{(\iota-\eta)}\lambda_0\,(\lambda_0/\lambda_1)^{\lambda_0/\Delta\lambda} \\ &\qquad\times\left(\sum_{\gamma=0}^{\eta}\left(\frac{\Delta\lambda}{r}\right)^{\eta-\gamma}\frac{\eta!}{\gamma!}\,\left(\ln\left[(\lambda_0/\lambda_1)^{\gamma}\,\Omega(p)\right]\right)^{\gamma}\right)(1-p)\,\Omega(p)^{\lambda_0/\Delta\lambda} \\ &+ C_n^{(\iota+1)}(1-p)\,\Omega(p)^{(r+\lambda_0)/\Delta\lambda}\,, \end{aligned}$$

where

$$\begin{aligned} d_1^{(\iota+1)}(k) &= \tfrac{\lambda_1}{r+\lambda_1}\left[\lambda_1 h + \tfrac{r}{\lambda_1}(\lambda_1 h - s)\,\iota\right] + \tfrac{r}{r+\lambda_1}\left[k\lambda_1 h + (1-k)s\right]\,, \\ d_0^{(\iota+1)}(k) &= \tfrac{\lambda_0}{r+\lambda_0}\left[\lambda_0 h - \tfrac{r}{\lambda_0}(s - \lambda_0 h)\,\iota\right] + \tfrac{r}{r+\lambda_0}\left[k\lambda_0 h + (1-k)s\right]\,, \\ m^{(\iota+1)} &= C^{(0)}\left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu}}{\lambda_0-\mu\Delta\lambda}\right)^{\iota}\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu}}{r+\lambda_0-\mu\Delta\lambda}\,, \\ \text{and}\quad l^{(\iota-\eta)} &= \frac{C^{(\iota-\eta)}}{\eta!}\left(-\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\lambda_0/\Delta\lambda}}{\Delta\lambda}\right)^{\eta} \quad \text{for } \eta = 0,\dots,\iota-1. \end{aligned}$$

This leads to the representation stated in the corollary. ■

Proof of Proposition 6: Let $\hat{p}_{N,N-1}$ denote the belief where the graph of V_N^* cuts $\mathcal{D}_{N-1}$, and consider $I = [\hat{p}_{N,N-1}-\epsilon, p^m]$ with $\epsilon > 0$ small enough that for all $\hat{p} \in I$, the unique solution to (1) that crosses $\mathcal{D}_{N-1}$ at the belief $\hat{p}$ has positive slope there.

Step 1: Construction of the average payoff function. Fix a belief $\hat{p} \in I$. On $[\hat{p}, 1]$, we define $u_{\hat{p}}$ as the unique solution to (1) that starts on $\mathcal{D}_{N-1}$ at $\hat{p}$. Starting from this initial condition, we then proceed iteratively as in the proof of Proposition 3, solving 'forward' towards lower beliefs and eventually to p_N^*. Between $\mathcal{D}_{N-1}$ and $\bar{\mathcal{D}}\wedge\mathcal{D}_{1-1/N}$, we solve the indifference ODDE (6); below $\bar{\mathcal{D}}\wedge\mathcal{D}_{1-1/N}$, we solve the ODDE

$$\Delta\lambda\, p(1-p)u'(p) = \lambda(p)[u(j(p)) - u(p)] - r[u(p)-s] + \tfrac{r}{N}[\lambda(p)h - s].$$

In this manner, we obtain a continuous function $u_{\hat{p}}$ on $[p_N^*, 1]$ such that: (i) $u_{\hat{p}}(p) = s + Nb(p, u_{\hat{p}}) - c(p)$ and $b(p, u_{\hat{p}}) > c(p)$ on $]\hat{p}, 1]$; (ii) $b(p, u_{\hat{p}}) = c(p)$ at all beliefs $p \in]p_N^*, \hat{p}]$

where the point $(p, u_{\hat{p}}(p))$ lies on or below $\mathcal{D}_{N-1}$ and above $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$; (iii) $u_{\hat{p}}(p) = s + b(p, u_{\hat{p}}) - c(p)/N$ and $b(p, u_{\hat{p}}) \leq c(p)$ at all beliefs $p \in]p_N^*, \hat{p}]$ where $(p, u_{\hat{p}}(p))$ lies on or below $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$.

Again proceeding as in the proof of Proposition 3, one establishes the existence of a $\hat{p} \in]\hat{p}_{N,N-1}, p^m[$ such that the corresponding function $u_{\hat{p}}$ has an interior global minimum equal to s at some belief $\breve{p} \in]p_N^*, \hat{p}[$. As $u'_{\hat{p}}(\breve{p}) = 0$, we have $\lambda(\breve{p})\,[u_{\hat{p}}(j(\breve{p})) - s]/r = b(\breve{p}, u_{\hat{p}}) = c(\breve{p})/N < c(\breve{p})$. For $\hat{p} = p^m$, on the other hand, the corresponding function $u_{\hat{p}}$ assumes value s at p^m. As its slope there is positive and $c(p^m) = 0$, we have $\lambda(p^m)\,[u_{\hat{p}}(j(p^m)) - s]/r > b(p^m, u_{\hat{p}}) = c(p^m)/N = c(p^m)$. By continuity of $u_{\hat{p}}$ with respect to $\hat{p}$, there exists $p_N^\ddagger \in]\hat{p}_{N,N-1}, p^m[$ such that $u_{p_N^\ddagger}$, the function $u_{\hat{p}}$ obtained for $\hat{p} = p_N^\ddagger$, has the following property: there is a belief $p_N^\flat \in]\breve{p}, p_N^\ddagger[$ such that $u_{p_N^\ddagger}(p_N^\flat) = s$, $u_{p_N^\ddagger}(p) > s$ for $p > p_N^\flat$, and $\lambda(p_N^\flat)\,[u_{p_N^\ddagger}(j(p_N^\flat)) - s]/r = c(p_N^\flat)$.

We define a function $\bar{u}$ on $[0,1]$ by taking $\bar{u} = u_{p_N^\ddagger}$ on $[p_N^\flat, 1]$, and $\bar{u} = s$ everywhere else. We want to establish that $\bar{u}$ is increasing on $[p_N^\flat, 1]$. The explicit representation (3) makes this obvious on $[p_N^\ddagger, 1]$. Moreover, the argument given in the proof of Proposition 3 shows that $\bar{u}$ is also increasing on $[p_N^\sharp, p_N^\ddagger]$ where $p_N^\sharp$ is the rightmost belief at which the graph of $\bar{u}$ crosses $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$. Suppose now that $\bar{u}$ is not increasing on $[p_N^\flat, p_N^\sharp]$. Then there exist beliefs $p' < p''$ in $]p_N^\flat, p_N^\sharp]$ such that $\bar{u}'(p'-) \geq 0$, $\bar{u}'(p''-) \leq 0$, and $\bar{u}$ is weakly decreasing on $[p', p'']$. As $j(p'') > j(p') > j(p_N^\flat) \geq p_N^\sharp$, we have $\bar{u}(j(p'')) > \bar{u}(j(p'))$, hence $\bar{u}(j(p'')) - \bar{u}(p'') > \bar{u}(j(p')) - \bar{u}(p')$ and $b(p'', \bar{u}) > b(p', \bar{u})$. This implies $\bar{u}(p') = s + b(p', \bar{u}) - c(p')/N < s + b(p'', \bar{u}) - c(p'')/N = \bar{u}(p'')$ – a contradiction.

Monotonicity immediately implies that $\bar{u}$ is the average payoff function associated with the following intensity of experimentation: $K(p) = N$ for $p \geq p_N^\ddagger$; $K(p) = N\bar{u}(p)/[(N-1)c(p)] < N$ for $p_N^\sharp < p < p_N^\ddagger$; $K(p) = 1$ for $p_N^\flat < p \leq p_N^\sharp$; and $K(p) = 0$ for $p \leq p_N^\flat$. Using the explicit form of the relevant ODDE to the left and right of $p_N^\sharp$, respectively, we see that $\Delta\lambda\, p_N^\sharp(1-p_N^\sharp)[\bar{u}'(p_N^\sharp+) - \bar{u}'(p_N^\sharp-)] = r[\bar{u}(p_N^\sharp) - s - (1-\frac{1}{N})c(p_N^\sharp)]$, so $\bar{u}$ has a kink at $p_N^\sharp$ with $\bar{u}'(p_N^\sharp-) > \bar{u}'(p_N^\sharp+)$ if and only if the intersection with $\bar{\mathcal{D}} \wedge \mathcal{D}_{1-1/N}$ is below $\mathcal{D}_{1-1/N}$; this kink then corresponds to a jump in the intensity of experimentation with $K(p_N^\sharp-) = 1 > K(p_N^\sharp+)$. By construction, K always jumps at $p_N^\flat$, and $\bar{u}$ always has a kink there. At all other beliefs, K is continuous, and $\bar{u}$ once continuously differentiable.

Step 2: Construction of the players' payoff functions and strategies. We define

$$\bar{b}(p, u) = [\lambda(p)\,(\bar{u}(j(p)) - u(p)) - \Delta\lambda\, p(1-p)u'(p)]/r$$

for any left-differentiable real-valued function u on $]0,1]$. (This is the benefit of experimentation when the value after a success is given by the payoff function $\bar{u}$.)

Fix any two beliefs $p_\ell < p_r$ in $[p_N^\flat, p_N^\sharp]$ and consider the four functions $u_{\ell F}, u_{\ell E}, u_{rF}$ and u_{rE} on $[p_\ell, p_r]$ that are uniquely determined by the following properties: $u_{\ell F}(p_\ell) = u_{\ell E}(p_\ell) = \bar{u}(p_\ell)$; $u_{rF}(p_r) = u_{rE}(p_r) = \bar{u}(p_r)$; on $]p_\ell, p_r]$, $u_{\ell F}$ and u_{rF} solve the free-rider ODE $u(p) = s + \bar{b}(p, u)$, while $u_{\ell E}$ and u_{rE} solve the experimenter ODE $u(p) = s + \bar{b}(p, u) - c(p)$. By construction, $[(N-1)u_{\ell F}+u_{\ell E}]/N$ coincides with $\bar{u}$ at p_ℓ and solves the same ODE as $\bar{u}$ on $]p_\ell, p_r]$, namely $u(p) = s + \bar{b}(p,u) - c(p)/N$, so it must coincide with $\bar{u}$ on $[p_\ell, p_r]$. The same argument applies to $[(N-1)u_{rF}+u_{rE}]/N$. We can thus conclude that $(N-1)u_{\ell F}+u_{\ell E} = (N-1)u_{rF}+u_{rE}$ on $[p_\ell, p_r]$.

Next, we have $u'_{\ell F}(p_\ell+) > \bar{u}'(p_\ell+)$ since $\lim_{p\downarrow p_\ell} \bar{b}(p,u_{\ell F}) = \bar{b}(p_\ell,\bar{u}) - c(p_\ell)/N < \bar{b}(p_\ell,\bar{u})$ and $\lim_{p\downarrow p_\ell}[\bar{u}(j(p)) - u_{\ell F}(p)] = \bar{u}(j(p_\ell)) - \bar{u}(p_\ell)$. Thus, $u_{\ell F}(p) > \bar{u}(p)$ immediately to the right of p_ℓ. Now, there cannot exist a belief $p' \in]p_\ell, p_r]$ such that $u_{\ell F}(p') = \bar{u}(p')$ and $u'_{\ell F}(p') \le \bar{u}'(p')$, because we would then have $c(p')/N = b(p',\bar{u}) - \bar{b}(p',u_{\ell F}) = -\Delta\lambda\, p'(1-p')[\bar{u}'(p') - u'_{\ell F}(p')]/r \le 0$ – a contradiction. This implies that $u_{\ell F} > \bar{u}$ on the entire interval $]p_\ell, p_r]$. Analogous arguments establish that $u_{\ell E} < \bar{u}$ on $]p_\ell, p_r]$ as well as $u_{rF} < \bar{u}$ and $u_{rE} > \bar{u}$ on $[p_\ell, p_r[$.

In particular, there exists a belief $p \in]p_\ell, p_r[$ such that $u_{\ell E}(p) = u_{rF}(p)$. Let p_1 denote the lowest such belief and define a continuous function $u_1(\cdot \mid p_\ell, p_r)$ on $[p_\ell, p_r]$ by setting $u_1(\cdot \mid p_\ell, p_r) = u_{\ell E}$ on $[p_\ell, p_1]$ and $u_1(\cdot \mid p_\ell, p_r) = u_{rF}$ on $[p_1, p_r]$. Using the identity $(N-1)u_{\ell F}+u_{\ell E} = (N-1)u_{rF}+u_{rE}$, we see that $u_{rE}(p_1) - u_{\ell F}(p_1) = (N-2)[u_{\ell F}(p_1) - u_{rF}(p_1)]$. If $N=2$, we define a continuous function $u_2(\cdot \mid p_\ell, p_r)$ on $[p_\ell, p_r]$ by setting $u_2(\cdot|p_\ell, p_r) = u_{\ell F}$ on $[p_\ell, p_1]$ and $u_2(\cdot|p_\ell, p_r) = u_{rE}$ on $[p_1, p_r]$.

If $N > 2$, we consider the function $u_{[2]}$ that coincides with $u_{\ell F}$ at p_1 and solves the experimenter ODE $u(p) = s + \bar{b}(p,u) - c(p)$ on $]p_1, p_r]$. As $u_{rE}(p_1) > u_{\ell F}(p_1)$, there is a belief $p \in]p_1, p_r[$ such that $u_{[2]}(p) = u_{rF}(p)$. Let p_2 denote the lowest such belief and define a continuous function $u_2(\cdot \mid p_\ell, p_r)$ on $[p_\ell, p_r]$ by setting $u_2(\cdot|p_\ell, p_r) = u_{\ell F}$ on $[p_\ell, p_1]$, $u_2(\cdot \mid p_\ell, p_r) = u_{[2]}$ on $[p_1, p_2]$ and $u_2(\cdot \mid p_\ell, p_r) = u_{rF}$ on $[p_2, p_r]$. By the same argument as above, $(N-2)u_{\ell F} + u_{[2]} + u_{rF} = (N-1)u_{rF} + u_{rE}$ on $[p_1, p_2]$, which is easily seen to imply $u_{rE}(p_2) - u_{\ell F}(p_2) = (N-3)[u_{\ell F}(p_2) - u_{rF}(p_2)]$. If $N=3$, we define a continuous function $u_3(\cdot \mid p_\ell, p_r)$ on $[p_\ell, p_r]$ by setting $u_3(\cdot \mid p_\ell, p_r) = u_{\ell F}$ on $[p_\ell, p_2]$ and $u_3(\cdot \mid p_\ell, p_r) = u_{rE}$ on $[p_2, p_r]$.

If $N > 3$, we proceed as in the previous paragraph to determine a belief $p_3 \in]p_2, p_r[$ and a continuous function $u_3(\cdot|p_\ell, p_r)$ that coincides with $u_{\ell F}$ on $[p_\ell, p_2]$, solves the experimenter ODE on $]p_2, p_3]$ and coincides with u_{rF} on $[p_3, p_r]$. Performing as many steps as necessary, we end up with beliefs $p_0 = p_\ell < p_1 < p_2 < \ldots < p_{N-1} < p_N = p_r$ and continuous functions $u_1(\cdot|p_\ell, p_r), \ldots, u_N(\cdot|p_\ell, p_r)$ on $[p_\ell, p_r]$ such that $u_n(\cdot|p_\ell, p_r)$ coincides with $u_{\ell F}$ on $[p_\ell, p_{n-1}]$, solves the experimenter ODE on $]p_{n-1}, p_n]$ and coincides with u_{rF} on $[p_n, p_r]$. By construction, the average of these N functions coincides with $\bar{u}$.

Now consider a finite family of contiguous intervals $]p_{\ell,i}, p_{r,i}]$ whose union equals $]p_N^\flat, p_N^\sharp]$. For each of these intervals, let $p_{n,i}$ denote the corresponding belief p_n as determined in the previous paragraph. Define functions $u_1, \ldots, u_N$ on the unit interval by setting $u_n = s$ on $[0, p_N^\flat]$, $u_n = u_n(\cdot \mid p_{\ell,i}, p_{r,i})$ on $]p_{\ell,i}, p_{r,i}]$ and $u_n = \bar{u}$ on $]p_N^\sharp, 1]$. For $n = 1, \ldots, N$, define a strategy k_n as follows: $k_n(p) = 1$ if p lies in $]p_N^\ddagger, 1]$ or one of the intervals $]p_{n-1,i}, p_{n,i}]$; $k_n(p) = 0$ if p lies in $[0, p_N^\flat]$ or one of the intervals $]p_{\ell,i}, p_{n-1,i}]$ and $]p_{n,i}, p_{r,i}]$; $k_n(p) = \bar{u}(p)/[(N-1)c(p)]$ everywhere else. Clearly, u_n is player n's payoff function associated with the strategy profile $(k_1, \ldots, k_N)$, and $\bar{u}$ is the corresponding average payoff function. Note that by construction, u_1 is differentiable at $p_N^\flat$ with $u_1'(p_N^\flat) = 0$.

Step 3: Ensuring mutually best responses. The following arguments establish that player n plays a best response against $K_{\neg n}$ as implied by the strategy profile constructed in Step 2. First, the graph of u_n is above $\mathcal{D}_{N-1}$ on $]p_N^\ddagger, 1]$, so playing R is optimal against $K_{\neg n}(p) = N - 1$ there. Second, $b(p, u_n) = c(p)$ on $]p_N^\sharp, p_N^\ddagger]$, making $k_n(p) = \bar{u}(p)/[(N-1)c(p)]$ trivially optimal on this interval. Third, after increasing the number and reducing the size of the intervals $]p_{\ell,i}, p_{r,i}]$ if necessary, the graph of u_n is above level s and below $\mathcal{D}_1$ on $]p_N^\flat, p_N^\sharp[$, so it is optimal for player n to play R whenever all other players play S, and to play S whenever one other player plays R. Fourth, we have $b(p_N^\flat, u_n) = \lambda(p_N^\flat)\,[u_n(j(p_N^\flat)) - s]/r = \lambda(p_N^\flat)\,[\bar{u}(j(p_N^\flat)) - s]/r = c(p_N^\flat)$ as the left-hand derivative of u_n at $p_N^\flat$ is zero, so playing S is optimal at this belief. Fifth, $u_n(j(p))$ is at least weakly increasing and $c(p)$ decreasing on $[0, p_N^\flat[$, therefore $b(p, u_n) < c(p)$ on this interval, again implying optimality of S. ■

Proof of Proposition 8: The proof proceeds in four steps, two of which are simpler versions of the corresponding steps in the proof of Proposition 6. Let $\hat{p}_{2,1}$ denote the belief where the graph of V_2^* cuts $\mathcal{D}_1$, and consider $I = [\hat{p}_{2,1} - \epsilon, p^m]$ with $\epsilon > 0$ small enough that for all $\hat{p} \in I$, the unique solution to (1) with $N = 2$ that crosses $\mathcal{D}_1$ at the belief $\hat{p}$ has positive slope there.

Step 1: Construction of the average payoff function. Fix a belief $\hat{p} \in I$. On $[\hat{p}, 1]$, we define $u_{\hat{p}}$ as the unique solution to (1) with $N = 2$ that starts on $\mathcal{D}_1$ at $\hat{p}$. On $[p_2^*, \hat{p}[$, we define $u_{\hat{p}}$ as the unique solution to the ODE

$$\Delta\lambda\, p(1-p)u'(p) + [r + \lambda(p)]u(p) = \tfrac{r}{2}[s + \lambda(p)h] + \lambda(p)u_{\hat{p}}(j(p))$$

that ends on $\mathcal{D}_1$ at $\hat{p}$. By construction, $u_{\hat{p}}$ is continuous, $u_{\hat{p}}(p) = s + 2b(p, u_{\hat{p}}) - c(p)$ on $]\hat{p}, 1]$, and $u_{\hat{p}}(p) = s + b(p, u_{\hat{p}}) - \frac{1}{2}c(p)$ on $]p_2^*, \hat{p}]$.

Proceeding as in the proof of Proposition 6, one establishes the existence of a $\hat{p} \in]\hat{p}_{2,1}, p^m[$ such that the corresponding function $u_{\hat{p}}$ has the following property: there is a belief $\bar{p} \in]p_2^*, \hat{p}[$ such that $u_{\hat{p}}(\bar{p}) = s$, $u_{\hat{p}}(p) > s$ for $p > \bar{p}$, and $\lambda(\bar{p})\,[u_{\hat{p}}(j(\bar{p})) - s]/r = c(\bar{p})$.

We define a function $\bar{u}$ on $[0, 1]$ by taking $\bar{u}$ equal to the function $u_{\hat{p}}$ just determined on $[\bar{p}, 1]$, and $\bar{u} = s$ everywhere else. We want to establish that $\bar{u}$ is increasing on $[\bar{p}, 1]$. The explicit representation (3) makes this obvious on $[\hat{p}, 1]$. Suppose now that $\bar{u}$ is not increasing on $[\bar{p}, \hat{p}]$. Then there exist beliefs $p' < p''$ in $]\bar{p}, \hat{p}]$ such that $\bar{u}'(p'-) \geq 0$, $\bar{u}'(p''-) \leq 0$, and $\bar{u}$ is weakly decreasing on $[p', p'']$. As $j(p'') > j(p') > p^m$, we have $\bar{u}(j(p'')) > \bar{u}(j(p'))$, hence $\bar{u}(j(p'')) - \bar{u}(p'') > \bar{u}(j(p')) - \bar{u}(p')$ and $b(p'', \bar{u}) > b(p', \bar{u})$. This implies $\bar{u}(p') = s + b(p', \bar{u}) - c(p')/2 < s + b(p'', \bar{u}) - c(p'')/2 = \bar{u}(p'')$ – a contradiction.

Step 2: Construction of the players' payoff functions and strategies. We define

$$\bar{b}(p, u) = [\lambda(p)\,(\bar{u}(j(p)) - u(p)) - \Delta\lambda\, p(1-p)u'(p)]/r$$

for any left-differentiable real-valued function u on $]0,1]$.

For any two beliefs $p_\ell < p_r$ in $[\bar{p},\hat{p}]$, Step 2 in the proof of Proposition 6 yields a belief $p_s \in\,]p_\ell, p_r[$ (denoted by p_1 there) as well as continuous functions $u_1(\cdot\,|\,p_\ell,p_r)$ and $u_2(\cdot\,|\,p_\ell,p_r)$ on $[p_\ell,p_r]$. Both functions coincide with $\bar{u}$ at the beliefs p_ℓ and p_r; $u_1(\cdot \mid p_\ell,p_r)$ solves $u(p) = s + \bar{b}(p,u) - c(p)$ on $]p_\ell,p_s]$ and $u(p) = s + \bar{b}(p,u)$ on $]p_s,p_r]$; $u_2(\cdot \mid p_\ell,p_r)$ solves $u(p) = s+\bar{b}(p,u)$ on $]p_\ell,p_s]$ and $u(p) = s+\bar{b}(p,u) - c(p)$ on $]p_s,p_r]$; the average of these two functions coincides with $\bar{u}$ on all of $[p_\ell,p_r]$.

It is easily seen that $u_1(\cdot\,|\,p_\ell,p_r) < \bar{u} < u_2(\cdot\,|\,p_\ell,p_r)$ on $]p_\ell,p_r[$. Moreover, using similar arguments as for the average payoff function, it is straightforward to show that $u_1(\cdot\,|\,p_\ell,p_r)$ is increasing on $[p_\ell,p_r]$, and $u_2(\cdot\,|\,p_\ell,p_r)$ on $[p_\ell,p_s]$. However, these arguments do not preclude the possibility that the function u_2 is decreasing on some interval $[p_s, p_s+\epsilon]$.

Now consider a finite family of contiguous intervals $]p_{\ell,i},p_{r,i}]$ whose union equals $]\bar{p},\hat{p}]$. For each of these intervals, let $p_{s,i}$ denote the corresponding belief p_s as determined above. Define functions u_1 and u_2 on the unit interval by setting $u_n = s$ on $[0,\bar{p}]$, $u_n = u_n(\cdot\,|\,p_{\ell,i},p_{r,i})$ on $]p_{\ell,i},p_{r,i}]$, and $u_n = \bar{u}$ on $]\hat{p},1]$. Define a simple strategy k_1 by setting $k_1(p) = 1$ if and only if p lies in $]\hat{p},1]$ or one of the intervals $]p_{\ell,i},p_{s,i}]$, and a simple strategy k_2 by setting $k_2(p) = 1$ if and only if p lies in $]\hat{p},1]$ or one of the intervals $]p_{s,i},p_{r,i}]$. Clearly, u_1 and u_2 are the payoff functions associated with the strategies k_1 and k_2, and $\bar{u}$ is the corresponding average payoff function. By construction, u_1 is differentiable at $\bar{p}$ with $u_1'(\bar{p}) = 0$.

Step 3: Establishing that $u_2'(\hat{p}) > -\Delta\lambda h$. Unlike u_1, the function u_2 is not necessarily increasing on $[\bar{p},\hat{p}]$, so we do not know whether its graph lies below the diagonal $\mathcal{D}_1$ to the left of $\hat{p}$, which will be important to establish the mutual best-response property in Step 4 below. Our next aim, therefore, is to show that $u_2'(\hat{p}) > -\Delta\lambda h$, implying that u_2 stays below $\mathcal{D}_1$ to the immediate left of $\hat{p}$.

We have $u_2(\hat{p}) = s+\bar{b}(\hat{p},u_2) - c(\hat{p}) = s + c(\hat{p})$, hence $\bar{b}(\hat{p},u_2) = 2c(\hat{p})$ and $\Delta\lambda\hat{p}\,(1-\hat{p})u_2'(\hat{p}) = \lambda(\hat{p})[\bar{u}(j(\hat{p})) - s - c(\hat{p})] - 2rc(\hat{p})$. As $\bar{u}(p) = \lambda(p)h + 2c(\hat{p})\frac{1-p}{1-\hat{p}}\left(\frac{\Omega(p)}{\Omega(\hat{p})}\right)^{\mu_2}$ on $[\hat{p},1]$, equation (2) for $N=2$ implies $\lambda(\hat{p})\bar{u}(j(\hat{p})) = \lambda_1^2 h\hat{p} + \lambda_0^2 h(1-\hat{p}) + 2\left[\frac{r}{2} + \lambda_0 - \mu_2\Delta\lambda\right]c(\hat{p})$. Straightforward computations now reveal that

$$u_2'(\hat{p}) = \Delta\lambda h - 2\left(\mu_2 + \frac{r}{2\Delta\lambda} + \hat{p}\right)\frac{c(\hat{p})}{\hat{p}\,(1-\hat{p})}$$

and that $u_2'(\hat{p}) > -\Delta\lambda h$ if and only if $\hat{p}$ is larger than

$$p^+ = \frac{(s-\lambda_0 h)(\mu_2 + \frac{r}{2\Delta\lambda})}{\lambda_1 h - s + \Delta\lambda h\,(\mu_2 + \frac{r}{2\Delta\lambda})}\,,$$

which is easily seen to lie between p_2^* and p^m. As $\hat{p} > \hat{p}_{2,1}$ (the belief where the graph of V_2^* cuts $\mathcal{D}_1$), a sufficient condition for $u_2'(\hat{p}) > -\Delta\lambda h$ is that $\hat{p}_{2,1} \geq p^+$ or, equivalently, $V_2^*(p^+) \leq s + c(p^+)$. Further computations show that the latter is the case if and only if

$$\frac{\mu_2 + \frac{r}{2\Delta\lambda} + 1}{\mu_2 + 1} \left(\frac{\mu_2}{\mu_2 + 1} \frac{\mu_2 + \frac{r}{2\Delta\lambda} + 1}{\mu_2 + \frac{r}{2\Delta\lambda}} \right)^{\mu_2} \leq 2.$$

Now, since $\frac{r}{2\Delta\lambda} < \mu_2$ and the function $h(x) = (\mu_2 + x + 1)^{\mu_2+1}(\mu_2 + x)^{-\mu_2}$ is increasing for $x \geq 0$, the left-hand side of this inequality is bounded above by

$$\frac{2\mu_2 + 1}{\mu_2 + 1} \left(\frac{\mu_2}{\mu_2 + 1} \frac{2\mu_2 + 1}{2\mu_2} \right)^{\mu_2},$$

which is clearly smaller than 2.

Step 4: Ensuring mutually best responses. As u_1 is increasing on $[\bar{p}, 1]$, player 1 is easily seen to play a best response against k_2, irrespective of the choice of intervals $[p_{\ell,i}, p_{r,i}]$. First, u_1 is above $\mathcal{D}_1$ on $]\hat{p}, 1]$. Second, it is above s and below $\mathcal{D}_1$ on $]\bar{p}, \hat{p}[$. Third, we have $b(\bar{p}, u_1) = \lambda(\bar{p})\,[u_1(j(\bar{p})) - s]/r = \lambda(\bar{p})\,[\bar{u}(j(\bar{p})) - s]/r = c(\bar{p})$ as the left-hand derivative of u_1 at $\bar{p}$ is zero. Fourth, $u_1(j(p))$ is at least weakly increasing and $c(p)$ decreasing on $[0, \bar{p}[$, therefore $b(p, u_1) < c(p)$ on this interval.

Turning to player 2, the fact that $u_2'(\hat{p}) > -\Delta\lambda h$ allows us to choose a finite family of intervals $[p_{\ell,i}, p_{r,i}]$ for any $\delta > 0$ such that the graph of u_2 is below the diagonal $\mathcal{D}_1$ on $]\bar{p}, \hat{p}[$ and the vertical distance $u_2 - u_1$ is at most δ at any belief in this interval (and hence on $[0, 1]$). If we take δ sufficiently small, player 2 is now also seen to play a best response. On $[\bar{p}, 1]$, the arguments are exactly the same as for player 1. On $[j^{-1}(\hat{p}), \bar{p}[$, $u_2(j(p)) = \bar{u}(j(p))$ is increasing and $c(p)$ decreasing, hence $b(p, u_2) < c(p)$. On $[0, j^{-1}(\hat{p})[$, finally, $u_2(j(p)) \leq u_1(j(p)) + \delta$ and $b(p, u_1) \leq b(j^{-1}(\hat{p}), u_1) < c(j^{-1}(\hat{p}))$, hence $b(p, u_2) \leq b(p, u_1) + \lambda(p)\,\delta/r < c(j^{-1}(\hat{p})) < c(p)$ for δ sufficiently small. ■

PROOF OF COROLLARY 4: The proof parallels that of Corollary 1. Here we consider the general case where K of the N players are playing R, for which the relevant ODE is

$$\Delta\lambda\, p(1-p)u'(p) + [\tfrac{r}{K} + \lambda(p)]u(p) = k\tfrac{r}{K}\lambda(p)h + (1-k)\tfrac{r}{K}s + \lambda(p)u^{(0)}(j(p))$$

with $k = 0$ for a free-rider and $k = 1$ for an experimenter. As in the proof of Corollary 1, we take $u^{(0)}(p) = \lambda(p)h + C^{(0)}\,(1-p)\,\Omega(p)^{\mu_N}$ with $C^{(0)}$ being the constant that fixes payoffs above the diagonal where everyone plays R. Having noted that under condition (7) the recursion ends after just one iteration, i.e. with $u^{(1)}$, we obtain equations for u_n of the form

$$\begin{aligned} u_n(p) \;=\; & \left(\lambda_1 h \frac{\lambda_1}{rK^{-1} + \lambda_1} + \frac{rK^{-1}}{rK^{-1} + \lambda_1} [k_n(p)\lambda_1 h + (1 - k_n(p))s] \right) p \\ & + \left(\lambda_0 h \frac{\lambda_0}{rK^{-1} + \lambda_0} + \frac{rK^{-1}}{rK^{-1} + \lambda_0} [k_n(p)\lambda_0 h + (1 - k_n(p))s] \right) (1-p) \\ & + C^{(0)} \left(\frac{\lambda_0\,(\lambda_0/\lambda_1)^{\mu_N}}{rK^{-1} + \lambda_0 - \mu_N\Delta\lambda} \right) (1-p)\,\Omega(p)^{\mu_N} + C_n^{(1)}(1-p)\,\Omega(p)^{(rK^{-1}+\lambda_0)/\Delta\lambda}, \end{aligned}$$

and setting $K = 1$, $N = 2$ leads to the representations stated in the corollary. ■

References

Bellman, R. and K.L. Cooke (1963): *Differential-Difference Equations.* New York: Academic Press.

Bergemann, D. and J. Välimäki (2008): "Bandit Problems," in *The New Palgrave Dictionary of Economics*, 2nd edition, editors S. Durlauf and L. Blume. Basingstoke and New York: Palgrave Macmillan Ltd.

Besanko, D. and J. Wu (2008): "The Impact of Market Structure and Learning on the Tradeoff between R&D Competition and Cooperations," working paper, Northwestern University.

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

Bolton, P. and C. Harris (2000): "Strategic Experimentation: the Undiscounted Case," in *Incentives, Organizations and Public Economics – Papers in Honour of Sir James Mirrlees*, editors P.J. Hammond and G.D. Myles. Oxford: Oxford University Press.

Cohen, A. and E. Solan (2008): "One-Arm Lévy Bandits," working paper, Tel Aviv University.

Davis, M.H.A. (1993): *Markov Models and Optimization.* London: Chapman & Hall.

Décamps, J.-P. and T. Mariotti (2004): "Investment Timing and Learning Externalities," *Journal of Economic Theory*, **118**, 80–102.

Hopenhayn, H. and F. Squintani (2008): "Preemption Games with Private Information," working paper, UCLA and University of Essex.

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

Klein, N. (2009): "Free-Riding and Delegation In Research Teams," working paper, University of Munich.

Klein, N. and S. Rady (2008): "Negatively Correlated Bandits," working paper, University of Munich.

Malueg, D.A. and S.O. Tsutsui (1997): "Dynamic R&D Competition with Learning," *RAND Journal of Economics*, **28**, 751–772.

Moscarini, G. and F. Squintani (2007): "Competitive Experimentation with Private Information," working paper, Yale University and University of Essex.

Presman, E.L. (1990): "Poisson Version of the Two-Armed Bandit Problem with Discounting," *Theory of Probability and its Applications*, **35**, 307–317.

Presman, E.L. and I.N. Sonin (1990): *Sequential Control with Incomplete Information: The Bayesian Approach to Multi-Armed Bandit Problems.* London: Academic Press.

Rothschild, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, **9**, 185–202.

Strulovici, B. (2008): "Learning While Voting: Determinants of Collective Experimentation," working paper, Northwestern University.