Strongly Symmetric Equilibria in Bandit Games*

Johannes Hörner† Nicolas Klein‡ Sven Rady§

This version: February 17, 2014

Abstract

This paper studies strongly symmetric equilibria (SSE) in continuous-time games of strategic experimentation with Poisson bandits. SSE payoffs can be studied via two functional equations similar to the HJB equation used for Markov equilibria that they generalize. This is valuable for three reasons. First, these equations retain the tractability of Markov equilibrium, while allowing for punishments and rewards: the best and worst equilibrium payoff are explicitly solved for. Second, they capture behavior of the discrete-time game: as period length goes to zero, the SSE payoff set converges to their solution. Third, they encompass a large payoff set: there is no perfect Bayesian equilibrium in the discrete-time game with frequent interactions achieving higher efficiency.

Keywords: Two-Armed Bandit, Bayesian Learning, Strategic Experimentation, Strongly Symmetric Equilibrium.

JEL Classification Numbers: C73, D83.

---

*We thank seminar participants at Aalto University Helsinki, Frankfurt, Paris (Séminaire Roy and Séminaire Parisien de Théorie des Jeux), Toulouse, the 2012 International Conference on Game Theory at Stony Brook, and the 2013 North American Summer Meeting of the Econometric Society for their comments and suggestions. Part of this paper was written during a visit to the Hausdorff Research Institute for Mathematics at the University of Bonn under the auspices of the Trimester Program "Stochastic Dynamics in Economics and Finance". Financial support from the Cowles Foundation and Deutsche Forschungsgemeinschaft (SFB/TR 15) is gratefully acknowledged.

†Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, johannes.horner@yale.edu.

‡Université de Montréal and CIREQ. Mailing address: Université de Montréal, Département de Sciences Économiques, C.P. 6128 succursale Centre-ville; Montréal, H3C 3J7, Canada, kleinnic@yahoo.com.

§University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany, rady@hcm.uni-bonn.de.

1 Introduction

There is a troubling disconnect between discrete-time and continuous-time game theory. With few exceptions, games in discrete time use either subgame-perfect equilibrium or, if there is incomplete information, perfect Bayesian equilibrium as a solution concept. With few exceptions, games in continuous time are concerned with Markov equilibria only. The technical reasons for this divide are well-known: defining outcomes, strategies and equilibrium in continuous time raises serious mathematical difficulties; restricting attention to Markov strategies bypasses these. Conceptually, however, the discontinuity is artificial and deeply unsatisfactory.

This paper proposes a middle ground. As we show, strongly symmetric equilibria retain the tractability of Markov equilibria. Markov perfect equilibrium payoffs can be studied via a well-known functional equation, the Hamilton-Jacobi-Bellman (or Isaacs) equation. Similarly, the set of strongly symmetric equilibrium payoffs is characterized by a pair of coupled functional equations. At the same time, unlike Markov equilibrium, strongly symmetric equilibrium allows for patterns of behavior that are both experimentally compelling and theoretically fundamental: punishments and rewards.

We confine our analysis to a particular class of models, the so-called two-armed bandit model, which has been extensively studied both in discrete and in continuous time (see, in particular, Keller et al. (2005) and Keller and Rady (2010)). More specifically, the set-up is as in Keller and Rady (2010). The motivation for this restriction is two-fold. First, the characterization of the appropriate boundary (or transversality) conditions for strongly symmetric equilibria hinges on fine details of the set-up, an analysis that we only know how to carry out within the confines of a specific model, as is also the case for Markov perfect equilibria. Second, restricting attention to such a well-studied model allows us to provide a closed-form for the equilibrium payoff set, a concrete illustration of how a slight weakening of the solution concept dramatically alters our understanding of incentives and expands the set of achievable payoffs.

Strongly symmetric equilibria (or SSE) are not new. They have been introduced to repeated games at least since Abreu (1986). They are known to be restrictive. To begin with, they make no sense if the model itself fails to be symmetric. But as Abreu (1986) already observes for repeated games, they are (i) easily calculated, being completely characterized by two simultaneous equations; (ii) more general than static Nash, or even Nash reversion, (iii) globally optimal in some cases, that is, for some parameters in his environment. See also Abreu, Pearce and Stacchetti (1986) for optimality of symmetric equilibria within a standard oligopoly framework, and Abreu, Pearce and Stacchetti (1993) for a motivation for the solution concept based on a notion of equal bargaining power. A more general analysis for repeated games with perfect monitoring is carried out by Cronshaw and Luenberger (1994) showing how the solution for the set of SSE payoffs is obtained by finding the largest scalar solving a certain equation.

These threes properties generalize to stochastic games (with "Markov perfect" replacing "Nash" in the three statements). Our first step involves establishing the rather straightforward functional analogues of the equations derived by Abreu, and Cronshaw and Luenberger for repeated games. This motivates the coupled functional equations in continuous-time that we put forth as a tool to analyze stochastic games as the bandit model. We then provide a formal limiting result to support the use of these functional equations to define the set of equilibria in continuous-time: the set of strongly symmetric equilibrium payoffs of the discrete-time game converges to the solution of these functional equations.

This is the central result of this paper. To be sure, we can and do directly define the set of strongly symmetric equilibria in continuous time as the set of solutions to these functional equations. After all, this is what is usually done with Markov equilibria. Checking systematically convergence of the equilibrium payoffs as defined in the discrete-time game to those as defined in the continuous-time game would defeat the purpose of exploiting the tractability of continuous time, whether for Markov or strongly symmetric equilibria. But given that, to the best of our knowledge, this paper is the first attempt at studying these coupled equations in continuous-time games, we view it as useful and reassuring to check that they capture precisely the strategic elements of the discrete-time game with frequent interactions in this particular instance. This is by no means a foregone conclusion: there are well-known examples in which the continuous-time definition of Markov equilibrium yields a set of payoffs that does not coincide with the limit of the set of Markov equilibrium payoffs for the discrete-time approximation. In fact, one corollary of our analysis is that the infinite-switching equilibria in Keller et al. (2005) have no counterpart in discrete time, no matter how small the time interval between consecutive choices; see also Heidhues, Rady and Strack (2012).

While proving this limit result requires some care, actually solving the continuous-time equations is a straightforward exercise in the case of the bandit model. This is where the analytical convenience of continuous time comes into play, yielding simple and exact solutions that admit intuitive interpretations. The resulting equilibrium payoff correspondence is rich: the symmetric Markov equilibrium is neither the lowest nor the highest selection. In fact, we show that the restriction to SSE is without loss in terms of joint payoffs: there is no sequence of perfect Bayesian equilibria in the discrete-time game whose limit (as we take the length of the time intervals to zero) sum of payoffs or experimentation rates would be higher than in the best SSE. The same holds true regarding the worst SSE joint payoff, which equals the single-agent payoff.

Both the best and the worst equilibrium are of the cutoff type, in which players experiment if and only if the belief exceeds a certain threshold. This contrasts with the non-existence of such equilibria within the set of Markov equilibria (see Proposition 3 of Keller and Rady (2010)). Cutoffs and equilibrium payoffs vary continuously with the parameters, including in the limit as a single success fully reveals the arm. The cutoff in the best equilibrium (identified in Lemma 1) satisfies an intuitive property: the less informative a success, the lower the cutoff. To understand this, note that, if a success is perfectly informative, as in Keller et al. (2005), there is no scope for providing incentives after a success, as it is strictly dominant to play the risky arm afterwards. In that case and only in that case does the cutoff coincide with the single-agent cutoff. The less informative a success is, the more players can exploit the continuation play after a success to enforce discipline. Whether or not the first-best, cooperative solution can be achieved hinges on a simple comparison: does a success at the cooperative threshold take the posterior belief above or below the single-agent threshold? If informativeness does not take the posterior above this threshold, the cooperative solution can be implemented. Unlike for Markov equilibria, comparative statics regarding this cutoff and best payoff are straightforward: the lower the payoff on the safe arm, or the more patient players are, the lower the cutoff. The same holds true if the number of players increases.

2 The Model

The basic setup is that of Keller et al. (2005) and Keller and Rady (2010). Time $t \in [0, \infty)$ is continuous, and the discount rate is $r > 0$. There are $N \geq 1$ players, each facing the same two-armed bandit problem with one safe and one risky arm.

The safe arm S generates a known expected payoff $s > 0$ per unit of time. The risky arm R generates lump-sum payoffs that are independent draws from a time-invariant distribution on $\mathbb{R} \backslash \{0\}$ with a known mean $h > 0$. These lump sums arrive at the jump times of a standard Poisson process whose intensity depends on an unknown state of the world, $\theta \in \{0, 1\}$. If $\theta = 1$, the intensity is $\lambda_1 > 0$ for all players; if $\theta = 0$, the intensity is λ_0 for all players with $0 \leq \lambda_0 < \lambda_1$. These constants are again known to the players. Conditional on θ, the Poisson processes that drive the payoffs of the risky arm are independent across players.

We are interested in discrete-time versions of the experimentation game where players can only adjust their actions at the times $t = 0, \Delta, 2\Delta, \ldots$ for some fixed $\Delta > 0$. The expected discounted payoff increment from using S for the length of time Δ is $\int_0^\Delta r\, e^{-r\,t}\, s\, dt = (1-\delta)s$ with $\delta = e^{-r\,\Delta}$. Conditional on θ, the expected discounted payoff increment from using R is $\mathbb{E}\Big[\int_0^\Delta r\, e^{-r\,t}\, h\, dN_{\theta,t}\Big]$ where $N_{\theta,t}$ is a standard Poisson process with intensity λ_θ; as $N_{\theta,t} - \lambda_\theta t$ is a martingale, this simplifies to $\int_0^\Delta r\, e^{-r\,t}\, h\, \lambda_\theta\, dt = (1-\delta)\lambda_\theta h$. We assume that $\lambda_0 h < s < \lambda_1 h$, so each player prefers R to S if R is good ($\theta = 1$), and prefers S to R if R is bad ($\theta = 0$).

Players start with a common prior belief about θ. Thereafter, they observe each other's actions and outcomes, so they hold common posterior beliefs throughout time. With p denoting the subjective probability that $\theta = 1$, the expected discounted payoff increment from using R conditional on all available information is $(1-\delta)\lambda(p)h$ with $\lambda(p) = p\lambda_1 + (1-p)\lambda_0$. This exceeds the payoff increment from using S if and only if p exceeds the myopic cutoff belief

$$p^m = \frac{s - \lambda_0 h}{(\lambda_1 - \lambda_0)h}\,.$$

To derive the law of motion of beliefs, consider one of the intervals of length Δ on which the player's actions $(k_1, \ldots, k_N) \in \{0,1\}^N$ are fixed, with $k_n = 1$ indicating that player n uses R, and $k_n = 0$ indicating that she uses S. With $K = \sum_{n=1}^N k_n$ players using the risky arm, the probability in state θ of a total of $j = 0, 1, 2, \ldots$ lump sums during this time interval is $\frac{(K\lambda_\theta\Delta)^j}{j!} e^{-K\lambda_\theta\Delta}$ by the sum property of the Poisson distribution. Given the belief p held at the beginning of the interval, therefore, the probability assigned to J lump sums arriving within the length of time Δ is

$$\Lambda^\Delta_{J,K}(p) = \frac{K^J \Delta^J}{J!}\left[p\lambda_1^J \gamma_1^K + (1-p)\lambda_0^J \gamma_0^K\right]$$

with $\gamma_\theta = e^{-\lambda_\theta\Delta}$, and the corresponding posterior belief is

$$B^\Delta_{J,K}(p) = \frac{p\lambda_1^J \gamma_1^K}{p\lambda_1^J \gamma_1^K + (1-p)\lambda_0^J \gamma_0^K}.$$

For $K > 0$, the absence of a lump-sum payoff over the length of time Δ makes players more pessimistic: $B^\Delta_{0,K}(p) < p$ whenever $p > 0$. Throughout the paper, we shall assume Δ small enough that $\lambda_1\gamma_1^N > \lambda_0\gamma_0^N$. This guarantees that successes always make players more optimistic: $B^\Delta_{J,K}(p) > p$ for all $J \geq 1$, $K > 0$ and $p < 1$.

For any bounded function w on $[0,1]$ and any $K \in \{0, 1, \ldots, N\}$, we define a bounded function $\mathcal{E}_K^\Delta w$ by

$$\mathcal{E}_K^\Delta w(p) = \sum_{J=0}^{\infty} \Lambda_{J,K}^\Delta(p)\, w(B_{J,K}^\Delta(p)).$$

This is the expectation of w with respect to the distribution of posterior beliefs when the current belief is p and K players use R for a length of time Δ.

A history of length $t = \Delta, 2\Delta, \ldots$ is a sequence

$$h_t = \left((k_{n,0})_{n=1}^N, (j_{n,\Delta})_{n=1}^N, \ldots, (k_{n,t-\Delta})_{n=1}^N, (j_{n,t})_{n=1}^N\right),$$

such that $k_{n,\tau} = 0 \Rightarrow j_{n,\tau+\Delta} = 0$. This history specifies all actions $k_{n,\tau} \in \{0,1\}$ taken by the players, and the resulting number of realized lump-sums $j_{n,\tau+\Delta} \in \mathbb{N}_0$. We write H_t for the set of all histories of length t, set $H_0 = \{\emptyset\}$, and let $H = \bigcup_{t=0,\Delta,2\Delta,\ldots}^\infty H_t$.

In addition, we assume that players have access to a public randomization device in every period, namely, a draw from the uniform distribution on $[0,1]$, which is assumed to be independent of θ and across periods. Following standard practice, we omit its realizations from the description of histories.

Along with the prior belief p_0, each profile of strategies induces a distribution over H. Given a history h_t, we can recursively define the sequence of beliefs p_τ through $p_\tau = B_{J_\tau, K_{\tau-\Delta}}^\Delta(p_{\tau-\Delta})$, where $J_\tau = \sum_{n=1}^N j_{n,\tau}$ and $K_{\tau-\Delta} = \sum_{n=1}^N k_{n,\tau-\Delta}$.[1]

A behavioral strategy σ_n for player n is a sequence $(\sigma_{n,\tau})_{\tau=0,\Delta,2\Delta,\ldots}$, where $\sigma_{n,\tau}$ is a map from H_τ to the set of probability distributions on $\{0,1\}$; a pure strategy takes values in the set of degenerate distributions only. A (pure or behavioral) strategy is a Markov (stationary) strategy if it depends on h_t only through the posterior belief p_t. It is symmetric if this map is the same for all players.

Player n seeks to maximize the average discounted expected payoff

$$(1-\delta)\,\mathbb{E}\left[\sum_{\ell=0}^{\infty} \delta^\ell \left\{\mathbb{1}\{k_{n,\ell\Delta} = 0\}s + \mathbb{1}\{k_{n,\ell\Delta} = 1\}\left[\mathbb{1}\{\theta = 1\}\lambda_1 + (1 - \mathbb{1}\{\theta = 1\})\,\lambda_0\right] h\right\}\right].$$

By the law of iterated expectations, this equals

$$(1-\delta)\,\mathbb{E}\left[\sum_{\ell=0}^{\infty} \delta^\ell \left\{\mathbb{1}\{k_{n,\ell\Delta} = 0\}s + \mathbb{1}\{k_{n,\ell\Delta} = 1\}\lambda(p_{\ell\Delta})h\right\}\right].$$

[1] Anticipating on the solution concept, this requires Bayes' rule to be applied off-path as well. This is a game of observable actions, so this raises no particular difficulty.

Nash equilibrium, perfect Bayesian equilibrium and Markov perfect equilibrium of the game with period length Δ are defined in the usual way.

Our focus is on pure-strategy strongly symmetric equilibria. As we shall see, the restriction to pure strategies entails no loss in terms of equilibrium payoffs, when $\lambda_0 > 0$ and we take the period length Δ to 0. A pure-strategy strongly symmetric equilibrium (SSE) is a perfect Bayesian equilibrium in which all players use the same pure strategy: $\sigma_n(h_t) = \sigma_{n'}(h_t)$, for all n, n' and $h_t \in H$. This implies symmetry of behavior after *any* history, not just on the equilibrium path of play. Note that any symmetric Markov perfect equilibrium is a strongly symmetric equilibrium. Endowing the set of histories with the product topology, the set of SSE is compact, and so is the set of SSE payoffs. If non-empty, this set is simply an interval in $\mathbb{R}$. Its characterization is the subject of the next section.

3 Characterizing Equilibrium Payoffs

Fix $\Delta > 0$. For $p \in [0,1]$, let $\overline{W}^{\Delta}(p)$ and $\underline{W}^{\Delta}(p)$ denote the supremum and infimum, respectively, of the set of payoffs over (pure-strategy) strongly symmetric equilibria, given prior belief p. If a pure-strategy SSE exists, these extrema are achieved, and $\overline{W}^{\Delta} \geq \underline{W}^{\Delta}$.

Proposition 1 *Suppose that $\overline{W}^{\Delta} \geq \underline{W}^{\Delta}$. The pair of functions $(\overline{w}, \underline{w}) = (\overline{W}^{\Delta}, \underline{W}^{\Delta})$ solve the functional equations*

$$\overline{w}(p) = \max_{\kappa \in \mathcal{K}(p;\overline{w},\underline{w})} \left\{ (1-\delta)[(1-\kappa)s + \kappa\lambda(p)h] + \delta \mathcal{E}^{\Delta}_{N\kappa} \overline{w}(p) \right\}, \tag{1}$$

$$\underline{w}(p) = \min_{\kappa \in \mathcal{K}(p;\overline{w},\underline{w})} \max_{k \in \{0,1\}} \left\{ (1-\delta)[(1-k)s + k\lambda(p)h] + \delta \mathcal{E}^{\Delta}_{(N-1)\kappa+k} \underline{w}(p) \right\}, \tag{2}$$

where $\mathcal{K}(p;\overline{w},\underline{w}) \subseteq \{0,1\}$ denotes the set of all κ such that

$$\begin{aligned} &(1-\delta)[(1-\kappa)s + \kappa\lambda(p)h] + \delta \mathcal{E}^{\Delta}_{N\kappa} \overline{w}(p) \\ &\quad \geq \max_{k \in \{0,1\}} \left\{ (1-\delta)[(1-k)s + k\lambda(p)h] + \delta \mathcal{E}^{\Delta}_{(N-1)\kappa+k} \underline{w}(p) \right\}. \end{aligned} \tag{3}$$

Moreover, $\underline{W}^{\Delta} \leq \underline{w} \leq \overline{w} \leq \overline{W}^{\Delta}$ for any solution $(\overline{w}, \underline{w})$ of (1)–(3).

PROOF: Fix a pair $(\overline{w}, \underline{w})$ that satisfies (1)–(3). Note that (1)–(2) imply that $\underline{w} \leq \overline{w}$. Given such a pair, and any prior p, we construct two SSE whose payoffs are respectively $\overline{w}$ and $\underline{w}$. It then follows that $\underline{W}^{\Delta} \leq \underline{w} \leq \overline{w} \leq \overline{W}^{\Delta}$. Let $\overline{\kappa}$ and $\underline{\kappa}$ denote a selection of the maximum and minimum of (1)–(2). The equilibrium strategies are described by a two-state automaton, whose states are referred to as "good" or "bad." The difference between the two equilibria lies in the initial state: $\overline{w}$ is achieved when the initial state is good, $\underline{w}$ when it is bad. In the good state, play proceeds according to $\overline{\kappa}$; in the bad state, according to $\underline{\kappa}$. Transitions are as follows. If the state is good and all players play $\overline{\kappa}$, play remains in the good state; otherwise, play shifts to the bad state. If after some history h, the state is bad and all players play $\underline{\kappa}$, play switches from the bad state to the good state with some probability $\eta(p) \in [0,1]$ where p is the belief held after history h. This switch is determined by the public randomization device (*i.e.*, the switch is a deterministic function of its realization). Otherwise, play remains in the bad state. The probability $\eta(p)$ is chosen so that

$$\begin{aligned} \underline{w}(p) \;=\; & (1-\delta)[(1-\underline{\kappa}(p))s + \underline{\kappa}(p)\lambda(p)h] \\ & + \delta\left\{\eta(p)\, \mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\overline{w}(p) + [1-\eta(p)]\, \mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\underline{w}(p)\right\}, \end{aligned} \tag{4}$$

with (1)–(3) ensuring that $\eta(p) \in [0,1]$. This completes the description of the strategies. The choice of η along with (1)–(2) rules out profitable one-shot deviations in either state, so that the automaton describes equilibrium strategies, and the desired payoffs are obtained.

It remains to show that $(\overline{W}^{\Delta}, \underline{W}^{\Delta})$ solve the functional equations whenever $\overline{W}^{\Delta} \geq \underline{W}^{\Delta}$. Note that in *any* SSE, given p, the action $\kappa(p)$ must be an element of $\mathcal{K}(p; \overline{W}^{\Delta}, \underline{W}^{\Delta})$. This is because the left-hand side of (3) with $\overline{w} = \overline{W}^{\Delta}$ is an upper bound on the continuation payoff if no player deviates, and the right-hand side with $\underline{w} = \underline{W}^{\Delta}$ a lower bound on the continuation payoff after a unilateral deviation. Consider the equilibrium that achieves $\overline{W}^{\Delta}$. Then

$$\overline{W}^{\Delta}(p) \leq \max_{\kappa \in \mathcal{K}(p; \overline{W}^{\Delta}, \underline{W}^{\Delta})} \left\{(1-\delta)[(1-\kappa)s + \kappa\lambda(p)h] + \delta\mathcal{E}^{\Delta}_{N\kappa}\overline{W}^{\Delta}(p)\right\},$$

as the action played must be in $\mathcal{K}(p; \overline{W}^{\Delta}, \underline{W}^{\Delta})$ and the continuation payoff is at most given by $\overline{W}^{\Delta}$. Similarly, $\underline{W}^{\Delta}$ must satisfy (2) with "$\geq$"instead of "=." Suppose now that the "$\leq$" were strict. Then we can define a strategy profile given prior p that (i) in period 0, plays the maximizer of the right-hand side, and (ii) from $t = \Delta$ onward, abides by the continuation strategy achieving $\overline{W}^{\Delta}(p_{\Delta})$. Because the initial action is in $\mathcal{K}(p; \overline{W}^{\Delta}, \underline{W}^{\Delta})$, this constitutes an equilibrium; and it achieves a payoff strictly larger than $\overline{W}^{\Delta}(p)$, a contradiction. Hence, (1) must hold with equality for $\overline{W}^{\Delta}$. The same reasoning applies to $\underline{W}^{\Delta}$ and (2). ■

4 Continuous Time

As Δ tends to 0, equations (1)–(2) transform into differential-difference equations involving terms that are familiar from Keller and Rady (2010). A formal Taylor approximation shows that for any $\kappa \in \{0,1\}$ and $K \in \{0,1,\ldots,N\}$,

$$\begin{aligned}&(1-\delta)[(1-\kappa)s+\kappa\lambda(p)h]+\delta\mathcal{E}_K^\Delta w(p)\\ &\quad = \quad w(p)+r\left\{(1-\kappa)s+\kappa\lambda(p)h+K\,b(p,w)-w(p)\right\}\Delta+o(\Delta)\end{aligned}$$

where

$$b(p,w)=\frac{\lambda(p)}{r}\,[w(j(p))-w(p)]-\frac{\lambda_1-\lambda_0}{r}\,p(1-p)\,w'(p)$$

and

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

As in Keller and Rady (2010), we can interpret $b(p,w)$ as the expected benefit of playing R when continuation payoffs are given by the function w. It weighs a discrete improvement in the overall payoff after a success against a marginal decrease in the absence of a success.[2]

Applying this approximation to (1)–(2), cancelling the terms of order 0 in Δ, dividing through by Δ, letting $\Delta \to 0$ and using the notation

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

for the opportunity cost of playing R, we obtain

$$\overline{w}(p) \quad = \quad s+\max_{\kappa\in\mathcal{K}(p;\overline{w},\underline{w})}\kappa\,[Nb(p,\overline{w})-c(p)]\,, \tag{5}$$

$$\underline{w}(p) \quad = \quad s+\min_{\kappa\in\mathcal{K}(p;\overline{w},\underline{w})}(N-1)\kappa\,b(p,\underline{w})+\max_{k\in\{0,1\}}k\,[b(p,\underline{w})-c(p)]\,. \tag{6}$$

To determine $\mathcal{K}(p;\overline{w},\underline{w})$ in the limit as $\Delta \to 0$, we will in general require Taylor expansions of higher order. We say that the action κ satisfies the incentive-compatibility constraint of order ℓ at p if we can write the difference between the left-hand and the right-hand sides of the inequality (3) as $a_{\kappa,\ell}(p)\Delta^\ell+o(\Delta^\ell)$ with a positive leading coefficient $a_{\kappa,\ell}(p)>0$. Thus, the incentive constraint of order 0 for $\kappa \in \{0,1\}$ is tantamount to the strict inequality $\overline{w}(p)>\underline{w}(p)$. The constraint of order 1 requires $\overline{w}(p)=\underline{w}(p)$

[2]As the belief is updated downward in the absence of a success, we can compute $b(p,w)$ whenever w possesses a left-hand derivative at p. In the following, we assume that the functions under consideration have left-hand derivatives of arbitrarily high order at each belief in the open unit interval.

as well as $Nb(p,\overline{w}) - c(p) > (N-1)b(p,\underline{w})$ (for $\kappa = 1$) or $c(p) > b(p,\underline{w})$ (for $\kappa = 0$), and so on. We say that κ satisfies the incentive-compatibility constraint of order ∞ at p if there is a $\bar{\Delta} > 0$ such that the two sides of (3) coincide for all $\Delta \in (0,\bar{\Delta})$. This is obviously the case for $\kappa = 0$ at $p = 0$, and for $\kappa = 1$ at $p = 1$.

We tentatively take $\mathcal{K}(p;\overline{w},\underline{w})$ to be the set of all $\kappa \in \{0,1\}$ that satisfy the incentive-compatibility constraint of order ℓ at p for some $\ell \in \{0,1,2,\ldots,\infty\}$. With this definition, any action κ that is incentive compatible at a belief p admits a $\Delta_{\kappa,p} > 0$ such that (3) holds for all $\Delta \in (0,\Delta_{\kappa,p})$. We will see below that this is actually not enough to capture all the restrictions that discrete-time incentive compatibility imposes in the limit as $\Delta \to 0$. At the boundaries of the unit interval, this will be no issue, of course, as we trivially have $\mathcal{K}(0;\overline{w},\underline{w}) = \{0\}$ and $\mathcal{K}(1;\overline{w},\underline{w}) = \{1\}$.

It is instructive to consider the unconstrained versions of (5)–(6), obtained by replacing $\mathcal{K}(p;\overline{w},\underline{w})$ with $\{0,1\}$ everywhere. The unconstrained version of (5) is

$$\overline{w}(p) = s + \max_{\kappa \in \{0,1\}} \kappa \left[Nb(p,\overline{w}) - c(p) \right]. \tag{7}$$

Its unique solution is the N-player cooperative value function in continuous time, denoted V_N^* in Keller and Rady (2010). It satisfies $V_N^*(p) = s$ for $p \leq p_N^*$, and $V_N^*(p) > s$ for $p > p_N^*$, where

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

and $\mu_N > 0$ is implicitly defined by

$$\frac{r}{N} + \lambda_0 - \mu_N \left(\lambda_1 - \lambda_0 \right) = \lambda_0 \left(\frac{\lambda_0}{\lambda_1} \right)^{\mu_N}.$$

On $(p_N^*, 1]$, we have

$$V_N^*(p) = \lambda(p)h + \frac{c(p_N^*)}{u(p_N^*;\mu_N)}\, u(p;\mu_N)$$

with

$$u(p;\mu) = (1-p) \left(\frac{1-p}{p} \right)^{\mu}.$$

V_N^* is once continuously differentiable, so that $Nb(p, V_N^*) - c(p)$ is continuous in p. This difference has a single zero at p_N^*, being positive to the right of it and negative to the left.

Setting $N = 1$, we obtain the single-agent value function V_1^* and corresponding cutoff p_1^*. This function is the unique solution of the unconstrained version of (6),

$$\underline{w}(p) = s + \min_{\kappa\in\{0,1\}} (N-1)\kappa\, b(p, \underline{w}) + \max_{k\in\{0,1\}} k\left[b(p,\underline{w}) - c(p)\right]. \tag{8}$$

In fact, as $b(p; V_1^*) \geq 0$ everywhere, we have $\min_{\kappa\in\{0,1\}}(N-1)\kappa\, b(p, V_1^*) = 0$, and (8) with this minimum set to zero is just the Bellman equation for V_1^*.

Returning to the constrained problem with its potentially smaller correspondence of feasible actions, we immediately see that a solution to (5)–(6) must satisfy $V_1^* \leq \underline{w} \leq \overline{w} \leq V_N^*$, and the infimum of the set of beliefs at which $\overline{w} > s$ must lie in $[p_N^*, p_1^*]$.

To locate this infimum more precisely, we consider a family of candidate solutions. For any $\underline{p}$ in the open unit interval interval, we define a continuous function $V_{N,\underline{p}}$ by setting

$$V_{N,\underline{p}}(p) = \lambda(p)h + \frac{c(\underline{p})}{u(\underline{p};\mu_N)}\, u(p;\mu_N)$$

for $p > \underline{p}$, and $V_{N,\underline{p}}(p) = s$ otherwise. From Keller and Rady (2010), we know that $V_{N,\underline{p}}$ is the players' common payoff function in continuous time when all N of them use the risky arm on $(\underline{p}, 1]$ and there is no experimentation otherwise; in particular, $V_{N,\underline{p}}(p) = s + Nb(p, V_{N,\underline{p}}) - c(p)$ on $(\underline{p}, 1]$. For $\underline{p} = p_N^*$, this is just the cooperative value function V_N^*. For $\underline{p} > p_N^*$, we have $V_{N,\underline{p}} < V_N^*$ on $(p_N^*, 1)$, and $V_{N,\underline{p}}$ is continuously differentiable except for a convex kink at $\underline{p}$, which implies a discontinuity in $Nb(p; V_{N,\underline{p}}) - c(p)$: this difference is positive on $(\underline{p}, 1]$, approaches zero as p tends to $\underline{p}$ from the right, is positive at $\underline{p}$ itself, and then decreases monotonically as p falls further, eventually assuming negative values.

Now consider the pair of functions $(\overline{w}, \underline{w}) = (V_{N,\underline{p}}, V_1^*)$ for a given $\underline{p} \in [p_N^*, p_1^*]$. On $(\underline{p}, 1)$, where $V_{N,\underline{p}} > V_1^*$, both actions are incentive compatible. On $(0, \underline{p}]$, where $V_{N,\underline{p}} = V_1^* = s$ and $V'_{N,\underline{p}} = (V_1^*)' = 0$ (with $V'_{N,\underline{p}}(\underline{p})$ meaning the left-hand derivative), $\kappa = 0$ is incentive compatible, satisfying the constraint of order 1 (if $\underline{p} < p_1^*$) or 2 (if $\underline{p} = p_1^*$) at $\underline{p}$, and the constraint of order 1 at lower beliefs. The incentive compatibility constraint of order 1 for $\kappa = 1$ on $(0, \underline{p}]$ hinges on how $Nb(p, V_{N,\underline{p}}) - c(p) = N\lambda(p)[V_{N,\underline{p}}(j(p)) - s]/r - c(p)$ compares with $(N-1)b(p, V_1^*) = (N-1)\lambda(p)[V_1^*(j(p)) - s]/r$, that is, on the sign of $\lambda(p)[NV_{N,\underline{p}}(j(p)) - (N-1)V_1^*(j(p)) - s] - rc(p)$. The situation at $p = \underline{p}$ is of particular interest.

Lemma 1 *There is a belief $\hat{p} \in [p_N^*, p_1^*]$ such that*

$$\lambda(\underline{p})\left[NV_{N,\underline{p}}(j(\underline{p})) - (N-1)V_1^*(j(\underline{p})) - s\right] - rc(\underline{p})$$

is negative if $0 < \underline{p} < \hat{p}$*, zero if* $\underline{p} = \hat{p}$*, and positive if* $\hat{p} < \underline{p} < 1$*. Moreover,* $\hat{p} = p_N^*$ *if and only if* $j(p_N^*) \leq p_1^*$*, and* $\hat{p} = p_1^*$ *if and only if* $\lambda_0 = 0$*.*

PROOF: See Appendix A. ■

A direct consequence of this result is that for $\underline{p} > \hat{p}$, $(\overline{w}, \underline{w}) = (V_{N,\underline{p}}, V_1^*)$ does not solve (5)–(6). In fact, $\underline{p} > \hat{p}$ implies that $\kappa = 1$ satisfies the incentive compatibility constraint of order 1 at $\underline{p}$. As $Nb(\underline{p}; V_{N,\underline{p}}) - c(\underline{p}) > 0$, this in turn means that $V_{N,\underline{p}}(\underline{p}) < s + \max_{\kappa \in \mathcal{K}(\underline{p}; V_{N,\underline{p}}, V_1^*)} \kappa \left[Nb(\underline{p}; V_{N,\underline{p}}) - c(\underline{p}) \right]$. In the case where $\hat{p} = p_N^*$, therefore, this means that $(\overline{w}, \underline{w}) = (V_{N,\underline{p}}, V_1^*)$ solves (5)–(6) if and only if $\underline{p} = p_N^*$ and hence $\overline{w} = V_N^*$.

Assume $\hat{p} > p_N^*$ now. For $p_N^* \leq \underline{p} < \hat{p}$, $(\overline{w}, \underline{w}) = (V_{N,\underline{p}}, V_1^*)$ solves (5)–(6) even though the leading coefficient in the Taylor expansion of the incentive-compatibility constraint for $\kappa = 1$ at $\underline{p}$ is negative, and hence $\mathcal{K}(\underline{p}; V_{N,\underline{p}}, V_1^*) = \{0\}$. (Recall that the constraint of order 0 is satisfied on $(\underline{p}, 1)$, and so $\mathcal{K}(p; V_{N,\underline{p}}, V_1^*) = \{0, 1\}$ there.)

This is the point where we have have to refine our definition of $\mathcal{K}(\cdot; \overline{w}, \underline{w})$ if, for fixed functions $\overline{w}$ and $\underline{w}$, we want to interpret the system (5)–(6) as the limit of (1)–(2) as $\Delta \to 0$. Note first that for every $\Delta > 0$, we can choose $p^\Delta > \underline{p}$ such that $p^\Delta \to \underline{p}$ and $V_{N,\underline{p}}(p^\Delta) - V_1^*(p^\Delta) = o(\Delta)$ as $\Delta \to 0$. This means that in the expansion of $\mathcal{E}_N^\Delta V_{N,\underline{p}}(p^\Delta) - \mathcal{E}_{N-1}^\Delta V_1^*(p^\Delta)$, it is the term of order 1 that dominates as Δ becomes small, not the term of order 0. Moreover, we can take p^Δ so that the posterior belief after $N-1$ or N failed experiments of length Δ is below $\underline{p}$, which implies that in the above expansion, the term of order 1 is determined entirely by the comparison of the "jump benefits" $N\lambda(p^\Delta)\Delta\,[V_{N,\underline{p}}(j(p^\Delta)) - s]$ and $(N-1)\lambda(p^\Delta)\Delta\,[V_1^*(j(p^\Delta)) - s]$. Therefore, the difference between the left-hand and the right-hand sides of (3) at p^Δ is Δ times $\lambda(p^\Delta)[NV_{N,\underline{p}}(j(p^\Delta)) - (N-1)V_1^*(j(p^\Delta)) - s] - rc(p^\Delta)$ plus terms of higher order. If $\underline{p} < \hat{p}$, this is negative for small Δ, implying that $\kappa = 1$ is not incentive compatible immediately to the right of $\underline{p}$.

We are thus led to the conclusion that the only appropriate solution to (5)–(6) is the pair of functions $(V_{N,\hat{p}}, V_1^*)$. This still begs the question, of course, whether this solution is the limit of the bounds $(\overline{W}^\Delta, \underline{W}^\Delta)$ on equilibrium payoffs in discrete time. In the following two sections, we will answer this question affirmatively.

Before doing so, we briefly compare the cutoff $\hat{p}$ with the belief at which all experimentation stops in the unique symmetric Markov perfect equilibrium of the continuous-time game.

Proposition 2 *For* $\lambda_0 > 0$*, the cutoff* $\hat{p}$ *is strictly lower than the belief at which all* *experimentation stops in the symmetric Markov perfect equilibrium of the continuous-time game.*

PROOF: Keller and Rady (2010) establish that in the unique symmetric Markov perfect equilibrium of the continuous-time game, all experimentation stops at the belief $\tilde{p}_N$ implicitly defined by $rc(\tilde{p}) = \lambda(\tilde{p}_N)[\tilde{u}(j(\tilde{p}_N)) - s]$, where $\tilde{u}$ is the players' common equilibrium payoff function. The results of Keller and Rady (2010) further imply that $V_{N,\tilde{p}_N}(j(\tilde{p}_N)) > \tilde{u}(j(\tilde{p}_N)) > V_1^*(j(\tilde{p}_N))$, so that $NV_{N,\tilde{p}_N}(j(\tilde{p}_N)) - (N-1)V_1^*(j(\tilde{p}_N)) > \tilde{u}(j(\tilde{p}_N))$, and hence $\hat{p} < \tilde{p}_N$ by Lemma 1. ■

The unique symmetric Markov perfect equilibrium in Keller and Rady (2010) implies a double-barrel inefficiency. Not only is the overall *amount* of experimentation too small, *i.e.* there is an inefficiently high probability of never finding out the true state of the world in the long run; the *speed* of experimentation is inefficiently slow to boot. The continuous-time limit of the strongly symmetric equilibria that we shall construct does better along both dimensions.

5 Bounds on Equilibrium Payoffs and the Range of Experimentation

For $\Delta > 0$, let $\tilde{p}^\Delta$ be the infimum of the set of beliefs at which the experimentation game with period length Δ admits a strongly symmetric equilibrium with payoff exceeding s. Let $\tilde{p} = \liminf_{\Delta \to 0} \tilde{p}^\Delta$. For small $\epsilon > 0$, consider the problem of maximizing the average of the players' payoffs in the discretized setting subject to symmetry of actions at all times and no use of R at beliefs $p \leq \tilde{p} - \epsilon$. Denote the corresponding value function by $\widetilde{W}^{\Delta,\epsilon}$. By definition of $\tilde{p}$, there exists a $\tilde{\Delta}_\epsilon > 0$ such that for $\Delta \in (0, \tilde{\Delta}_\epsilon)$, the function $\widetilde{W}^{\Delta,\epsilon}$ provides an upper bound on the players' common payoffs in any strongly symmetric equilibrium, and hence $\overline{W}^\Delta \leq \widetilde{W}^{\Delta,\epsilon}$. As the optimal solution to this constrained optimization problem is feasible for a similarly constrained planner in continuous time, we have $\widetilde{W}^{\Delta,\epsilon} \leq V_{N,p_\epsilon}$ with $p_\epsilon = \max\{\tilde{p} - \epsilon, p_N^*\}$. Lemma B.3 in the Appendix establishes that $\widetilde{W}^{\Delta,\epsilon} \to V_{N,p_\epsilon}$ uniformly as $\Delta \to 0$.

As any player can choose to ignore the information contained in the other players' experimentation results, the value function W_1^Δ of a single agent experimenting in isolation constitutes an obvious lower bound on a player's payoff in any (not just strongly symmetric) equilibrium, and so we have $\underline{W}^\Delta \geq W_1^\Delta$. Lemma B.4 (applied for $\bar{p} = 1$) establishes uniform convergence $W_1^\Delta \to V_1^*$ as $\Delta \to 0$.

Now, fix $\epsilon > 0$ and consider a sequence of Δ's smaller than $\tilde{\Delta}_\epsilon$ and converging to 0 such that the corresponding beliefs $\tilde{p}^\Delta$ converge to $\tilde{p}$. For each Δ in this sequence, choose $p^\Delta > \tilde{p}^\Delta$ such that $B^\Delta_{0,N-1}(p^\Delta) < \tilde{p}^\Delta$, and hence $B^\Delta_{0,N}(p^\Delta) < \tilde{p}^\Delta$ as well. If the players start at the belief p^Δ, therefore, and $N-1$ or all of them use R for Δ units of time without success, then the posterior belief ends up below $\tilde{p}^\Delta$ and there is no further experimentation in equilibrium. Now, playing R at p^Δ (against $N-1$ players who do so) yields at most

$$\begin{aligned}
&(1-\delta)\lambda(p^\Delta)h + \delta\left\{\Lambda^\Delta_{0,N}(p^\Delta)s + \sum_{J=1}^{\infty}\Lambda^\Delta_{J,N}(p^\Delta)\widetilde{W}^{\Delta,\epsilon}(B^\Delta_{J,N}(p^\Delta))\right\} \\
&= r\Delta\,\lambda(p^\Delta)h + (1-r\Delta)\left\{[1-N\lambda(p^\Delta)\Delta]s \right. \\
&\qquad \left. + N\lambda(p^\Delta)\Delta\,\widetilde{W}^{\Delta,\epsilon}(B^\Delta_{1,N}(p^\Delta))\right\} + o(\Delta) \\
&= s + \left\{r[\lambda(\tilde{p})h - s] + N\lambda(\tilde{p})[V_{N,p_\epsilon}(j(\tilde{p})) - s]\right\}\Delta + o(\Delta),
\end{aligned}$$

while playing S yields at least

$$\begin{aligned}
&(1-\delta)s + \delta\left\{\Lambda^\Delta_{0,N-1}(p^\Delta)s + \sum_{J=1}^{\infty}\Lambda^\Delta_{J,N-1}(p^\Delta)W^\Delta_1(B^\Delta_{J,N-1}(p^\Delta))\right\} \\
&= r\Delta\,s + (1-r\Delta)\left\{[1-(N-1)\lambda(p^\Delta)\Delta]s \right. \\
&\qquad \left. + (N-1)\lambda(p^\Delta)\Delta\,W^\Delta_1(B^\Delta_{1,N-1}(p^\Delta))\right\} + o(\Delta) \\
&= s + \left\{(N-1)\lambda(\tilde{p})[V^*_1(j(\tilde{p})) - s]\right\}\Delta + o(\Delta).
\end{aligned}$$

Incentive compatibility of R at p^Δ for small Δ requires

$$\lambda(\tilde{p})\left[NV_{N,p_\epsilon}(j(\tilde{p})) - (N-1)V^*_1(j(\tilde{p})) - s\right] - rc(\tilde{p}) \geq 0.$$

Letting $\epsilon \to 0$, we have $p_\epsilon \to \tilde{p}$ and thus

$$\lambda(\tilde{p})\left[NV_{N,\tilde{p}}(j(\tilde{p})) - (N-1)V^*_1(j(\tilde{p})) - s\right] - rc(\tilde{p}) \geq 0. \tag{9}$$

By Lemma 1, this means $\tilde{p} \geq \hat{p}$. For any $\epsilon > 0$ and $\Delta \in (0, \tilde{\Delta}_\epsilon)$, therefore, the set of beliefs at which experimentation can be sustained in a strongly symmetric equilibrium of the discrete-time game with period length Δ is contained in the interval $(\hat{p}-\epsilon, 1]$, and we have the chain of inequalities $\overline{W}^\Delta \leq \widetilde{W}^{\Delta,\epsilon} \leq V_{N,p_\epsilon} \leq V_{N,\hat{p}-\epsilon}$. Upon letting

$\epsilon \to 0$, this yields $\limsup_{\Delta \to 0} \overline{W}^{\Delta}(p) \le V_{N,\hat{p}}(p)$ for all p.

In the following section, we show constructively that these bounds on the range of experimentation and the best and worst equilibrium payoffs are tight, that is, $\tilde{p} = \hat{p}$ and, for all p, $\lim_{\Delta \to 0} \overline{W}^{\Delta}(p) = V_{N,\hat{p}}(p)$ and $\lim_{\Delta \to 0} \underline{W}^{\Delta}(p) = V_1^*(p)$.

6 Strongly Symmetric Equilibria for Small Δ

6.1 The Non-Revealing Case ($\lambda_0 > 0$)

The equilibrium construction for $\lambda_0 > 0$ is inspired by the first part of the proof of Proposition 1. For sufficiently small $\Delta > 0$, we shall exhibit a strongly symmetric equilibrium that can be summarized by two functions, $\overline{\kappa}$ and $\underline{\kappa}$, which will not depend on Δ. The equilibrium strategy is characterized by a two-state automaton. In the "good" state, play proceeds according to $\overline{\kappa}$ and the equilibrium payoff satisfies

$$\overline{w}^{\Delta}(p) = (1-\delta)[(1-\overline{\kappa}(p))s + \overline{\kappa}(p)\lambda(p)h] + \delta\mathcal{E}^{\Delta}_{N\overline{\kappa}(p)}\overline{w}^{\Delta}(p), \tag{10}$$

while in the "bad" state, play proceeds according to $\underline{\kappa}$ and the payoff satisfies

$$\underline{w}^{\Delta}(p) = \max_{k} \left\{(1-\delta)[(1-k)s + k\lambda(p)h] + \delta\mathcal{E}^{\Delta}_{(N-1)\underline{\kappa}(p)+k}\underline{w}^{\Delta}(p)\right\}. \tag{11}$$

That is, $\underline{w}^{\Delta}$ is the value from a player's best response to all other players following $\underline{\kappa}$.

A unilateral deviation from $\overline{\kappa}$ in the good state is punished by a transition to the bad state in the following period; otherwise we remain in the good state. If there is no unilateral deviation from $\underline{\kappa}$ in the bad state, a draw of a public randomization device determines whether the state next period is good or bad (and guarantees that the payoff is indeed given by $\underline{w}^{\Delta}$); otherwise we remain in the bad state.

With continuation payoffs given by $\overline{w}^{\Delta}$ and $\underline{w}^{\Delta}$, the common action $\kappa \in \{0,1\}$ can be sustained at a belief p if and only if

$$\begin{aligned}&(1-\delta)[(1-\kappa)s + \kappa\lambda(p)h] + \delta\mathcal{E}^{\Delta}_{N\kappa}\overline{w}^{\Delta}(p) \\ &\quad\ge\ (1-\delta)[\kappa s + (1-\kappa)\lambda(p)h] + \delta\mathcal{E}^{\Delta}_{(N-1)\kappa+1-\kappa}\underline{w}^{\Delta}(p).\end{aligned} \tag{12}$$

The functions $\overline{\kappa}$ and $\underline{\kappa}$ define an SSE, therefore, if and only if (12) holds for $\kappa = \overline{\kappa}(p)$ and $\kappa = \underline{\kappa}(p)$ at all p.

The probability $\eta^{\Delta}(p)$ of a transition from the bad to the good state in the absence of a unilateral deviation from $\underline{\kappa}(p)$ is then pinned down by the requirement that

$$\begin{aligned}\underline{w}^{\Delta}(p) &= (1-\delta)[(1-\underline{\kappa}(p))s+\underline{\kappa}(p)\lambda(p)h] \\ &\quad + \delta\left\{\eta^{\Delta}(p)\,\mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\overline{w}^{\Delta}(p)+[1-\eta^{\Delta}(p)]\,\mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\underline{w}^{\Delta}(p)\right\}.\end{aligned} \tag{13}$$

If $k=\underline{\kappa}(p)$ is optimal in (11), we simply set $\eta^{\Delta}(p)=0$. Otherwise, (11) and (12) imply

$$\delta\mathcal{E}^{\Delta}_{N\overline{\kappa}(p)}\overline{w}^{\Delta}(p)\geq\underline{w}^{\Delta}(p)-(1-\delta)[(1-\underline{\kappa}(p))s+\underline{\kappa}(p)\lambda(p)h]>\delta\mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\underline{w}^{\Delta}(p),$$

so (13) holds with

$$\eta^{\Delta}(p)=\frac{\underline{w}^{\Delta}(p)-(1-\delta)[(1-\underline{\kappa}(p))s+\underline{\kappa}(p)\lambda(p)h]-\delta\mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\underline{w}^{\Delta}(p)}{\delta\mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\overline{w}^{\Delta}(p)-\delta\mathcal{E}^{\Delta}_{N\underline{\kappa}(p)}\underline{w}^{\Delta}(p)}\in(0,1].$$

We specify the functions $\overline{\kappa}$ and $\underline{\kappa}$ as follows. Given $\underline{p}\in(\hat{p},p_1^*)$ and $\bar{p}\in(p^m,1)$, let

$$\overline{\kappa}(p)=\begin{cases}1 & \text{if } p>\underline{p},\\ 0 & \text{if } p\leq\underline{p},\end{cases}$$

while

$$\underline{\kappa}(p)=\begin{cases}1 & \text{if } p>\bar{p},\\ 0 & \text{if } p\leq\bar{p}.\end{cases}$$

Note that punishment and reward strategies agree outside of $(\underline{p},\bar{p})$. The continuous-time payoff function associated with the common Markov strategy $\overline{\kappa}$ is $V_{N,\underline{p}}$; we write $V_{1,\bar{p}}$ for the continuous-time payoff function obtained from a best response against the opponents' common strategy $\underline{\kappa}$. In Appendix B, we establish uniform convergence $\overline{w}^{\Delta}\to V_{N,\underline{p}}$ and $\underline{w}^{\Delta}\to V_{1,\bar{p}}$ as $\Delta\to 0$, and $V_{1,\bar{p}}\to V_1^*$ as $\bar{p}\to 1$.

Proposition 3 *For $\lambda_0>0$, there are beliefs $p^\flat\in(\hat{p},p_1^*)$ and $p^\sharp\in(p^m,1)$ such that for all $\underline{p}\in(\hat{p},p^\flat)$ and $\bar{p}\in(p^\sharp,1)$, there exists $\bar{\Delta}>0$ such that for all $\Delta\in(0,\bar{\Delta})$, the two-state automaton with functions $\overline{\kappa}$ and $\underline{\kappa}$ defines a strongly symmetric perfect Bayesian equilibrium of the experimentation game with period length Δ.*

PROOF: See Appendix A. ■

6.2 The Fully Revealing Case ($\lambda_0 = 0$)

As we have seen *supra*, as long as $\lambda_0 > 0$ and breakthroughs are not fully revealing, there is some scope for the provision of intertemporal incentives *after any history* such that the players' belief remains above $\hat{p}$. This is no longer the case for $\lambda_0 = 0$, though. This can be seen in Condition (9), which we can rewrite as follows:

$$rc(\tilde{p}) \leq \lambda(\tilde{p}) \left[V_{N,\tilde{p}}(j(\tilde{p})) - s + (N-1)[V_{N,\tilde{p}}(j(\tilde{p})) - V_1^*(j(\tilde{p}))] \right].$$

At beliefs such that a single failed round of experimentation takes us below $\tilde{p}$, intertemporal incentives can be provided conditional on there being a success in this potentially last round of experimentation, by virtue of the term $(N-1)[V_{N,\tilde{p}}(j(\tilde{p})) - V_1^*(j(\tilde{p}))]$. If $\lambda_0 > 0$, there exists a $\nu > 0$ such that, for all $\tilde{p} \in [\hat{p}, p^m]$, $(N-1)[V_{N,\tilde{p}}(j(\tilde{p})) - V_1^*(j(\tilde{p}))] > \nu$. This means that it is always possible to provide intertemporal incentives that are substantial; indeed, they are bounded away from 0, while an agent's benefit from deviating is of order Δ, since he can only enjoy the fruits of his deviation for a single period. Yet, if $\lambda_0 = 0$, $j(\tilde{p}) = 1$, and $V_{N,\tilde{p}}(j(\tilde{p})) = V_1^*(j(\tilde{p})) = \lambda_1 h$, implying that intertemporal incentives cannot be provided 'close to $\tilde{p}$', as now any success takes us to a belief of one, so that everyone will play risky forever in any equilibrium. This raises the possibility of unravelling. If we cannot support incentives just above the candidate threshold below which play proceeds according to the symmetric Markov equilibrium, will the actual threshold not "shoot up"?

To settle whether unravelling occurs or not requires us to study the discrete-time game in considerable detail. As already mentioned, we do not claim that the specific choice of the discrete-time game is innocuous: it might well be that requiring players to move in alternate periods, for instance, would yield different conclusions.

Because the optimality equations for the discrete-time game are less tractable than their continuous-time analogue, the details of their derivation and their properties are intricate and relegated to the Appendix.[3]

First, we show that there is no perfect Bayesian equilibrium with any experimentation at beliefs below the single-agent cutoff $p_1^\Delta = \inf\{p \colon W_1^\Delta(p) > s\}$.

---

[3]These difficulties are already present for the study of the Markov equilibrium in discrete time. Unlike in the continuous-time limit, in which an explicit solution is known (see Keller et al. (2005)), the symmetric MPE in discrete time does not seem to admit an easy solution. In fact, there are open sets of beliefs for which there are multiple symmetric Markov equilibria in discrete time, no matter how small Δ. It is not known whether these discrete-time equilibria all converge (in some sense) to the symmetric equilibrium of Keller et al. (2005); in fact, it is not known whether *some* discrete-time Markov equilibrium converges to it.

Lemma 2 *Let $\lambda_0 = 0$. Fix $\Delta > 0$ and any prior belief $p < p_1^\Delta$. Then the unique perfect Bayesian equilibrium outcome specifies that all players play safe in all periods.*[4]

Proof: See Appendix C. ■

Lemma 2 already rules out the possibility that the asymmetric equilibria of Keller et al. (2005) with an infinite number of switches can be approximated in discrete time. The highest payoff that can be hoped for, then, involves all players experimenting above p_1^Δ.

Unlike for the case $\lambda_0 > 0$ (see Proposition 3), an explicit description of a two-state automaton implementing strongly symmetric equilibria whose payoffs converge to the obvious upper and lower bounds appears elusive. This is because, for beliefs that are arbitrarily close to (but above) p_1^Δ, equilibrium strategies are necessarily mixed. The proof of the next proposition establishes that the length of the interval of beliefs for which this is the case is vanishing as $\Delta \to 0$. In particular, for higher beliefs (except for beliefs arbitrarily close to 1, when $\kappa(p) = 1$ is strictly dominant), both pure actions can be enforced in equilibrium.

Proposition 4 *For $\lambda_0 = 0$, there are beliefs $p^\S \in (p_1^*, p^m)$ and $p^\sharp \in (p^m, 1)$ such that for all $\underline{p} \in (p_1^*, p^\S)$ and $\bar{p} \in (p^\sharp, 1)$, there exists $\bar{\Delta} > 0$ such that for all $\Delta \in (0, \bar{\Delta})$, there exists*

- *a strongly symmetric equilibrium in which the equilibrium strategy specifies $\kappa(p) = 1$ for all $p > \underline{p}$ and $\kappa(p) = 0$ for all $p < p_1^*$ along the equilibrium path (starting from a prior $p^0 > \underline{p}$) ;*
- *a strongly symmetric equilibrium in which the players' payoff does not exceed their best-reply payoff against $\kappa(p) = 0$ for all $p \notin [p_1^*, \underline{p}] \cup [\bar{p}, 1]$, $\kappa(p) = 1$ otherwise on the equilibrium path (starting from $p^0 \in (\underline{p}, \bar{p})$).*

Proof: See Appendix C. ■

While this proposition is somewhat weaker than Proposition 3 –the analogous statement for the case $\lambda_0 > 0$– its implications for limit payoffs (as $\Delta \to 0$) are the same; indeed, given that the interval $[p_1^*, \underline{p}]$ can be chosen to be arbitrarily small (as the proof establishes, of the order Δ as it turns out), its impact on equilibrium payoffs

[4] This does not extend to off-path behavior, of course. If a player deviates by pulling the risky arm and obtains a success, players all switch to the risky arm from that point on.

starting from $p^0 > \underline{p}$ is of order Δ: that is for both strongly symmetric equilibria whose existence is stated in the Proposition, the payoff converges to the one in which, respectively, $\kappa(p) = 1$ for all $p \geq p_1^*$ (for the first SSE, given $p^0 > \underline{p}$) and the best-reply payoff against $\kappa(p) = 0$ for all $p > p_1^*$ (for the second SSE, given $p^0 \in (\underline{p}, \bar{p})$).

6.3 Limit SSE Payoffs

Recall that, for fixed Δ, we write $\overline{W}^\Delta$ and $\underline{W}^\Delta$ for the pointwise supremum and infimum, respectively, of the set of strongly symmetric equilibrium payoff functions. The main result of this section is a characterization of $\overline{W}^\Delta$ and $\underline{W}^\Delta$ as the period length vanishes.

Proposition 5 $\lim_{\Delta\to 0} \overline{W}^\Delta = V_{N,\hat{p}}$ *and* $\lim_{\Delta\to 0} \underline{W}^\Delta = V_1^*$, *uniformly on* $[0, 1]$.

PROOF: For $\lambda_0 > 0$ and a given $\epsilon > 0$, the explicit representation for $V_{N,\underline{p}}$ in Section 4 and the uniform convergence $V_{1,\bar{p}} \to V_1^*$ as $\bar{p} \to 1$ (established in Lemma B.5) allow us to choose $\xi > 0$, $\underline{p} \in (\hat{p}, p^\S)$ (for some $p^\S > \hat{p}$) and $\bar{p} \in (p^\sharp, 1)$ such that $\|V_{N,\hat{p}-\xi} - V_{N,\hat{p}}\| < \epsilon$, $\|V_{N,\underline{p}} - V_{N,\hat{p}}\| < \epsilon$ and $\|V_{1,\bar{p}} - V_1^*\| < \frac{\epsilon}{2}$, with $\|\cdot\|$ denoting the supremum norm. Note that the same holds for $\lambda_0 = 0$ (no reference to equilibrium was made). Next, for $\lambda_0 > 0$, Proposition 3, Lemma B.7, Section 5 and Lemma B.4 imply the existence of a $\Delta^\dagger > 0$ such that for all $\Delta \in (0, \Delta^\dagger)$, the two-state automaton defined by the cutoffs $\underline{p}$ and $\bar{p}$ constitutes an SSE of the game with period length Δ and the following inequalities hold: $\overline{w}^\Delta \geq V_{N,\underline{p}}$, $\overline{W}^\Delta \leq V_{N,\hat{p}-\xi}$, $\|\underline{w}^\Delta - V_{1,\bar{p}}\| < \frac{\epsilon}{2}$ and $\|W_1^\Delta - V_1^*\| < \epsilon$. For $\Delta \in (0, \Delta^\dagger)$, we thus have

$$V_{N,\hat{p}} - \epsilon < V_{N,\underline{p}} \leq \overline{w}^\Delta \leq \overline{W}^\Delta \leq V_{N,\hat{p}-\xi} < V_{N,\hat{p}} + \epsilon$$

and

$$V_1^* - \epsilon < W_1^\Delta \leq \underline{W}^\Delta \leq \underline{w}^\Delta < V_{1,\bar{p}} + \frac{\epsilon}{2} < V_1^* + \epsilon,$$

so that $\|\overline{W}^\Delta - V_{N,\hat{p}}\|$ and $\|\underline{W}^\Delta - V_1^*\|$ are both smaller than ϵ, which was to be shown.

For the case $\lambda_0 = 0$, the proof of Proposition 4 establishes that there exists $K \in I\!N$ such that, given $\underline{p}$ as stated in the Proposition, we can take $\bar{\Delta}$ as $(\underline{p} - p_1^*)/K$. Equivalently, $p_1^* + K\bar{\Delta} = \underline{p}$. Hence, Proposition 4 can be restated as saying that, for some $\bar{\Delta} > 0$, and all $\Delta \in (0, \bar{\Delta})$, there exists $p_\Delta \in (p_1^*, p_1^* + K\Delta)$ such that the two conclusions of the Proposition hold with $\underline{p} = p_\Delta$. Let $\overline{w}^\Delta, \underline{w}^\Delta$ denote the payoffs from the first and second SSE from the Proposition, respectively, fixing the prior.[5] Given that $\underline{p} \to p_1^*$ and $\overline{w}^\Delta(p) \to s$, $\underline{w}^\Delta(p) \to s$ for all $p \in (p_1^*, p_\Delta)$, it follows that we can pick $\Delta^\dagger \in (0, \bar{\Delta})$ such that for all $\Delta \in (0, \Delta^\dagger)$, $\overline{w}^\Delta \geq V_{N,\underline{p}} - \epsilon$, and as before, $\overline{W}^\Delta \leq V_{N,\hat{p}-\xi}$, $\|\underline{w}^\Delta - V_{1,\bar{p}}\| < \frac{\epsilon}{2}$ and $\|W_1^\Delta - V_1^*\| < \epsilon$. The obvious inequalities follow as before, subtracting an additional ϵ to the left-hand side of the first one; and the conclusion follows as before, using 2ϵ as an upper bound. ■

7 Perfect Bayesian Equilibrium Payoffs

Our next result shows that the restriction to strongly symmetric equilibria is innocuous.

Proposition 6 *In the limit as $\Delta \to 0$, perfect Bayesian equilibria generate the same range of experimentation and the same set of average payoffs per player as strongly symmetric equilibria.*

Proof: See Appendix A. ■

8 Comparative Statics

Our next result concerns the question as to what happens when the players become infinitely patient or impatient. If players are myopic, they will not react to future rewards and punishments. It is therefore no surprise that in this case the cooperative solution cannot be sustained in equilibrium. By contrast, if players are very patient, the planner's solution can be sustained in equilibrium provided the number of players is large enough.

Lemma 3 *For $\lambda_0 > 0$,*

$$\lim_{r \to 0} \frac{j(p_N^*)}{p_1^*} = \frac{\lambda_1}{N\lambda_0},$$

and

$$\lim_{r \to \infty} \frac{j(p_N^*)}{p_1^*} = \frac{\lambda_1 h}{s}.$$

Proof: See Appendix A. ■

[5] Hence, to be precise, these payoffs are only defined on those beliefs that can be reached given the prior and the equilibrium strategies.

In the following lemma, we show that, in the case $\lambda_0 > 0$, the more players participate in the game the more they will engage in experimentation.

Lemma 4 *For $\lambda_0 > 0$, $\hat{p}$ is decreasing in N.*

PROOF: See Appendix A. ■

Appendix

A Proofs

Proof of Lemma 1: We start by noting that given the functions V_1^* and V_N^*, the cutoffs p_1^* and p_N^* are uniquely determined by

$$\lambda(p_1^*)[V_1^*(j(p_1^*)) - s] = rc(p_1^*) \tag{A.1}$$

and

$$\lambda(p_N^*)[NV_N^*(j(p_N^*)) - Ns] = rc(p_N^*), \tag{A.2}$$

respectively.

Consider the differentiable function f on $(0, 1)$ given by

$$f(\underline{p}) = \lambda(\underline{p})[NV_{N,\underline{p}}(j(\underline{p})) - (N-1)V_1^*(j(\underline{p})) - s] - rc(\underline{p}).$$

For $\lambda_0 = 0$, we have $j(\underline{p}) = 1$ and $V_{N,\underline{p}}(j(\underline{p})) = V_1^*(j(\underline{p})) = \lambda_1 h$ for all p, so $f(\underline{p}) = \lambda(\underline{p})[V_1^*(j(\underline{p})) - s] - rc(\underline{p})$, which is zero at $\underline{p} = p_1^*$ by (A.1), positive for $\underline{p} > p_1^*$, and negative for $\underline{p} < p_1^*$.

Assume $\lambda_0 > 0$. For $0 < \underline{p} < p \leq 1$, we have $V_{N,\underline{p}}(p) = \lambda(p)h + c(\underline{p})u(p; \mu_N)/u(\underline{p}; \mu_N)$ with the function $u(p; \mu) = (1-p)\left(\frac{1-p}{p}\right)^{\mu}$ which is strictly convex for $\mu > 0$. Moreover, we have $V_1^*(p) = s$ when $p \leq p_1^*$, and $V_1^*(p) = \lambda(p)h + Cu(p; \mu_1)$ with a constant $C > 0$ otherwise. Using the fact that

$$u(j(p); \mu) = \frac{\lambda_0}{\lambda(p)}\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu} u(p; \mu),$$

we see that the term $\lambda(\underline{p})NV_{N,\underline{p}}(j(\underline{p}))$ is actually linear in $\underline{p}$. When $j(\underline{p}) \leq p_1^*$, the term $-\lambda(\underline{p})(N-1)V_1^*(j(\underline{p}))$ is also linear in $\underline{p}$; when $j(\underline{p}) > p_1^*$, the nonlinear part of this term simplifies to $-(N-1)C\lambda_0^{\mu_1+1}u(\underline{p}; \mu_1)/\lambda_1^{\mu_1}$. This shows that f is concave, and strictly concave on the interval of all $\underline{p}$ for which $j(\underline{p}) > p_1^*$. As $\lim_{\underline{p}\to 1} f(\underline{p}) > 0$, this in turn implies that f has at most one root in the open unit interval; if so, f assumes negative values to the left of the root, and positive values to the right.

As $V_{N,p_1^*}(j(p_1^*)) > V_1^*(j(p_1^*))$, moreover, we have

$$f(p_1^*) > \lambda(p_1^*)[V_1^*(j(p_1^*)) - s] - rc(p_1^*) = 0$$

by (A.1). The potential root of f must thus lie in $[0, p_1^*)$. If $j(p_N^*) \leq p_1^*$, then $V_1^*(j(p_N^*)) = s$ and

$$f(p_N^*) = \lambda(p_N^*)[NV_N^*(j(p_N^*)) - Ns] - rc(p_N^*) = 0$$

by (A.2). If $j(p_N^*) > p_1^*$, then $V_1^*(j(p_N^*)) > s$ and $f(p_N^*) < 0$, so f has a root in (p_N^*, p_1^*). ■

PROOF OF PROPOSITION 3: We take $p^\flat$ as in Lemma B.8; Lemma B.9 ensures that $p^\flat > \hat{p}$. We fix a $\underline{p} \in (\hat{p}, p^\flat)$. By Lemma 1,

$$\lambda(p)[NV_{N,p}(j(p)) - (N-1)V_1^*(j(p)) - s] - rc(p) > 0$$

on $[\underline{p}, 1]$. As $V_{N,p}(j(p)) \leq V_{N,\underline{p}}(j(p))$ for $p \geq \underline{p}$, this implies

$$\lambda(p)[NV_{N,\underline{p}}(j(p)) - (N-1)V_1^*(j(p)) - s] - rc(p) > 0$$

on $[\underline{p}, 1]$. By Lemma B.5, there exists a belief $p^\sharp > p^m$ such that for all $\bar{p} > p^\sharp$, $\inf\{p : V_{1,\bar{p}}(p) > s\} \in (\underline{p}, p_1^*)$ and

$$\lambda(p)[NV_{N,\underline{p}}(j(p)) - (N-1)V_{1,\bar{p}}(j(p)) - s] - rc(p) > 0 \tag{A.3}$$

on $[\underline{p}, 1]$. We fix a $\bar{p} \in (p^\sharp, 1)$ and define $p^\dagger = \inf\{p \colon V_{1,\bar{p}}(p) > s\}$.

By Lemmas B.7 and B.8, there is a $\Delta_0 > 0$ such that $\overline{w}^\Delta \geq V_{N,\underline{p}} \geq \underline{w}^\Delta$ on the unit interval for all $\Delta < \Delta_0$. For any such Δ and any $p \in [0, \underline{p}]$, the common action $\kappa = \overline{\kappa}(p) = \underline{\kappa}(p) = 0$ trivially satisfies the incentive constraint (12). In fact, since $\underline{w}^\Delta(p) = s$, we have $(1-\delta)s + \delta\mathcal{E}_0^\Delta \underline{w}^\Delta(p) \geq (1-\delta)\lambda(p)h + \delta\mathcal{E}_1^\Delta \underline{w}^\Delta(p)$ by (11); as $\overline{w}^\Delta \geq \underline{w}^\Delta$, this in turn implies $(1-\delta)s + \delta\mathcal{E}_0^\Delta \overline{w}^\Delta(p) \geq (1-\delta)\lambda(p)h + \delta\mathcal{E}_1^\Delta \underline{w}^\Delta(p)$.

For all $\Delta < \Delta_0$ and $p \in (\bar{p}, 1]$, moreover, the common action $\kappa = \overline{\kappa}(p) = \underline{\kappa}(p) = 1$ satisfies the incentive constraint (12) because $\lambda(p)h > s$ and $\mathcal{E}_N^\Delta \overline{w}^\Delta(p) \geq \mathcal{E}_N^\Delta V_{N,\underline{p}}(p) \geq \mathcal{E}_{N-1}^\Delta V_{N,\underline{p}}(p) \geq \mathcal{E}_{N-1}^\Delta \underline{w}^\Delta(p)$, where the second of these inequalities follows from convexity of $V_{N,\underline{p}}$.

Now, let $\nu_1 > 0$ be such that

$$\lambda(p)[NV_{N,\underline{p}}(j(p)) - (N-1)V_{1,\bar{p}}(j(p)) - s] - rc(p) > \nu_1 \tag{A.4}$$

for all $p \in [\underline{p}, \bar{p}]$. Such a ν_1 exists by (A.3) and the continuity of its left-hand side in p. Fix $p^\ddagger \in (\underline{p}, p^\dagger)$ such that

$$(N\lambda(p^\ddagger) + r)\left[V_{N,\underline{p}}(p^\ddagger) - s\right] < \nu_1/3. \tag{A.5}$$

By Lemma B.4, there exists a $\Delta_1 \in (0, \Delta_0)$ such that for $\Delta < \Delta_1$, $\underline{w}^\Delta(p) = s$ on $[0, p^\ddagger]$. By the same argument as above, this implies that for these Δ, the common action $\kappa = \underline{\kappa}(p) = 0$ satisfies the incentive constraint (12) on $(\underline{p}, p^\ddagger]$ as well.

In the remainder of the proof, we simplify the notation by writing p_j^K for $B_{j,K}^\Delta(p)$, the posterior belief starting from p when K players use the risky arm and j of them receive a lump-sum within the length of time Δ.

For $p \in (\underline{p}, p^\ddagger]$ and $\kappa = \overline{\kappa}(p) = 1$, the left-hand side of the incentive constraint (12) expands as

$$\begin{aligned} & r\Delta\,\lambda(p)h + (1-r\Delta)\left\{N\lambda(p)\Delta\,\overline{w}^\Delta(p_1^N) + (1-N\lambda(p)\Delta)\,\overline{w}^\Delta(p_0^N)\right\} + O(\Delta^2) \\ = \ & \overline{w}^\Delta(p_0^N) + \left\{r\lambda(p)h + N\lambda(p)\overline{w}^\Delta(p_1^N) - (N\lambda(p)+r)\overline{w}^\Delta(p_0^N)\right\}\Delta + O(\Delta^2), \end{aligned}$$

and the right-hand side as

$$
\begin{aligned}
& r\Delta\, s + (1-r\Delta)\left\{(N-1)\lambda(p)\Delta\, \underline{w}^\Delta(p_1^{N-1}) + [1-(N-1)\lambda(p)\Delta]\, \underline{w}^\Delta(p_0^{N-1})\right\} + O(\Delta^2) \\
= \;& \underline{w}^\Delta(p_0^{N-1}) + \left\{rs + (N-1)\lambda(p)\underline{w}^\Delta(p_1^{N-1}) - [(N-1)\lambda(p)+r]\underline{w}^\Delta(p_0^{N-1})\right\}\Delta + O(\Delta^2).
\end{aligned}
$$

For $\Delta < \Delta_1$, we have $\overline{w}^\Delta(p_0^N) \geq s = \underline{w}^\Delta(p_0^{N-1})$, so the difference between the left-hand and right-hand sides is no smaller than Δ times

$$
\lambda(p)\left[N\overline{w}^\Delta(p_1^N) - (N-1)\underline{w}^\Delta(p_1^{N-1}) - s\right] - rc(p) - (N\lambda(p)+r)\left[\overline{w}^\Delta(p_0^N) - s\right]
$$

plus terms of order Δ^2 and higher.

Let $\epsilon = \frac{\nu_1}{15(N\lambda_1+r)}$. By Lemmas B.6 and B.4 as well as Lipschitz continuity of $V_{N,\underline{p}}$ and $V_{1,\bar{p}}$, there exists $\Delta_2 \in (0,\Delta_1)$ such that for $\Delta < \Delta_2$, $\|\overline{w}^\Delta - V_{N,\underline{p}}\|$, $\|\underline{w}^\Delta - V_{1,\bar{p}}\|$, $\max_{\underline{p}\leq p\leq p^\ddagger} |V_{N,\underline{p}}(p_1^N) - V_{N,\underline{p}}(j(p))|$ and $\max_{\underline{p}\leq p\leq p^\ddagger} |V_{1,\bar{p}}(p_1^{N-1}) - V_{1,\bar{p}}(j(p))|$ are all smaller than ϵ. For $\Delta < \Delta_2$, we thus have

$$
\begin{aligned}
\overline{w}^\Delta(p_1^N) &> V_{N,\underline{p}}(j(p)) - 2\epsilon, \\
\underline{w}^\Delta(p_1^{N-1}) &< V_{1,\bar{p}}(j(p)) + 2\epsilon, \\
\overline{w}^\Delta(p_0^N) &< V_{N,\underline{p}}(p_0^N) + \epsilon,
\end{aligned}
$$

so that the expression displayed above is larger than $\nu_1 - [(5N-2)\lambda(p)+r]\epsilon - \nu_1/3 > \nu_1/3$ by (A.4), (A.5) and the definition of ϵ. This implies that there is a $\Delta_3 \in (0,\Delta_2)$ such that for all $\Delta < \Delta_3$, the incentive constraint (12) holds for $\overline{\kappa}$ on $(\underline{p}, p^\ddagger]$.

As $V_{N,\underline{p}} > V_{1,\bar{p}}$ on $(\underline{p}, 1)$, there exist $\Delta_4 \in (0,\Delta_3)$ and $\nu_2 > 0$ such that

$$
V_{N,\underline{p}}(p_0^{N-1}) - V_{1,\bar{p}}(p_0^{N-1}) > \nu_2 \tag{A.6}
$$

for all $\Delta < \Delta_4$ and $p \in (p^\ddagger, \bar{p}]$. At any belief p in this interval, the difference between the left-hand and right-hand sides of (12) for $\kappa = \overline{\kappa}(p) = 1$ is $\overline{w}^\Delta(p_0^N) - \underline{w}^\Delta(p_0^{N-1}) + O(\Delta)$. By Lemmas B.6 and B.4 and Lipschitz continuity of $V_{N,\underline{p}}$, there exists $\Delta_4 \in (0,\Delta_3)$ such that for $\Delta < \Delta_4$, $\|\overline{w}^\Delta - V_{N,\underline{p}}\|$, $\|\underline{w}^\Delta - V_{1,\bar{p}}\|$ and $\max_{p^\ddagger\leq p\leq\bar{p}} |V_{N,\underline{p}}(p_0^N) - V_{N,\underline{p}}(p_0^{N-1})|$ are all smaller than $\nu_2/4$. For $\Delta < \Delta_4$ and $p \in (p^\ddagger, \bar{p})$, we thus have $\overline{w}^\Delta(p_0^N) > V_{N,\underline{p}}(p_0^N) - \nu_2/4 > V_{N,\underline{p}}(p_0^{N-1}) - \nu_2/2$ and $\underline{w}^\Delta(p_0^{N-1}) < V_{1,\bar{p}}(p_0^{N-1}) + \nu_2/4$, so that by (A.6) the difference between the left-hand and right-hand sides of (12) for $\kappa = \overline{\kappa}(p) = 1$ is larger than $\nu_2/4 + O(\Delta)$. Thus, there is a $\Delta_5 \in (0,\Delta_4)$ such that for all $\Delta < \Delta_5$, (12) holds for $\overline{\kappa}$ on $(p^\ddagger, \bar{p}]$.

For $p \in (p^\ddagger, \bar{p}]$ and $\kappa = \underline{\kappa}(p) = 0$, the difference between the left-hand and right-hand sides of (12) is $\overline{w}^\Delta(p) - \underline{w}^\Delta(p_0^1) + O(\Delta)$, and the same steps as in the previous paragraph yield existence of a $\bar{\Delta} \in (0,\Delta_5)$ such that for all $\Delta < \bar{\Delta}$, the incentive constraint (12) for $\underline{\kappa}$ is also satisfied on $(p^\ddagger, \bar{p}]$. ■

Proof of Proposition 6: For any given $\Delta > 0$, let $\check{p}^\Delta$ be the infimum of the set of beliefs at which there is some (possibly asymmetric) perfect Bayesian equilibrium that gives a payoff

$w_n(p) > s$ to at least one player. Let $\breve{p} = \liminf_{\Delta\to 0} \breve{p}^\Delta$. By construction, $\breve{p} \leq \hat{p}$.

For any fixed $\epsilon > 0$ and $\Delta > 0$, consider the problem of maximizing the players' average payoff subject to no use of R at beliefs $p \leq \breve{p} - \epsilon$, and write $\breve{W}^{\Delta,\epsilon}$ for the corresponding value function. Let $\breve{p}_\epsilon = \max\{\breve{p} - \epsilon, p_N^*\}$. Uniform convergence $\breve{W}^{\Delta,\epsilon} \to V_{N,\breve{p}_\epsilon}$ follows from the same arguments as in the proof of Lemma B.3.

Consider a sequence of Δ's converging to 0 such that the corresponding beliefs $\breve{p}^\Delta$ converge to $\breve{p}$. For each Δ in this sequence, select a perfect Bayesian equilibrium as well as a belief $p^\Delta > \breve{p}^\Delta$ starting from which a single failed experiment takes us below $\breve{p}^\Delta$. Let L^Δ be the number of players who, at the initial belief p^Δ, play R with positive probability in the selected equilibrium. Let L be an accumulation point of the sequence of L^Δ's. After selecting a subsequence of Δ's, we can assume without loss of generality that player $n = 1, \ldots, L$ plays R with probability $\alpha_n^\Delta > 0$ at p^Δ, while player $n = L+1, \ldots, N$ plays S; we can further assume that $(\alpha_n^\Delta)_{n=1}^L$ converges to a limit $(\alpha_n)_{n=1}^L$ in $[0,1]^L$.

For player $n = 1, \ldots, L$ to play optimally at p^Δ, it must be the case that

$$\begin{aligned}
&(1-\delta)\left[\alpha_n^\Delta \lambda(p^\Delta)h + (1-\alpha_n^\Delta)s\right] + \delta\left\{\Pr^\Delta(\emptyset)w_{n,\emptyset}^\Delta + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)w_{n,I,J}^\Delta\right\} \\
&\quad\geq\ (1-\delta)s + \delta\left\{\Pr_{-n}^\Delta(\emptyset)w_{n,\emptyset}^\Delta + \sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)w_{n,I,J}^\Delta\right\},
\end{aligned}$$

where we write $\Pr^\Delta(I)$ for the probability that the set of players experimenting is $I \subseteq \{1, \ldots, L\}$, $\Pr_{-n}^\Delta(I)$ for the probability that among the $L-1$ players in $\{1, \cdots, L\} \setminus \{n\}$ the set of players experimenting is I, and $w_{n,I,J}^\Delta$ for the conditional expectation of player n's continuation payoff given that exactly the players in I were experimenting and had J successes ($w_{n,\emptyset}^\Delta$ is player n's continuation payoff if no one was experimenting). As $\Pr^\Delta(\emptyset) = (1-\alpha_n^\Delta)\Pr_{-n}^\Delta(\emptyset) \leq \Pr_{-n}^\Delta(\emptyset)$, the inequality continues to hold when we replace $w_{n,\emptyset}^\Delta$ by its lower bound s. After subtracting $(1-\delta)s$ from both sides, we then have

$$\begin{aligned}
&(1-\delta)\alpha_n^\Delta\left[\lambda(p^\Delta)h - s\right] + \delta\left\{\Pr^\Delta(\emptyset)s + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)w_{n,I,J}^\Delta\right\} \\
&\quad\geq\ \delta\left\{\Pr_{-n}^\Delta(\emptyset)s + \sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)w_{n,I,J}^\Delta\right\}.
\end{aligned}$$

Summing up these inequalities over $n = 1, \cdots, L$ and writing $\bar{\alpha}^\Delta = \frac{1}{L}\sum_{n=1}^L \alpha_n^\Delta$ yields

$$
\begin{aligned}
&(1-\delta)L\bar{\alpha}^\Delta\left[\lambda(p^\Delta)h - s\right] + \delta\left\{\Pr^\Delta(\emptyset)Ls + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)\sum_{n=1}^{L} w_{n,I,J}^\Delta\right\} \\
&\geq \delta\left\{\sum_{n=1}^{L}\Pr_{-n}^\Delta(\emptyset)s + \sum_{n=1}^{L}\sum_{K=1}^{L-1}\sum_{|I|=K,\, n\notin I}\Pr_{-n}^\Delta(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)w_{n,I,J}^\Delta\right\}.
\end{aligned}
$$

By construction, $w_{n,I,0}^\Delta = s$ whenever $I \neq \emptyset$. For $|I| = K > 0$ and $J > 0$, moreover, we have $w_{n,I,J}^\Delta \geq W_1^\Delta(B_{J,K}^\Delta(p^\Delta))$ for *all* players $n = 1, \ldots, N$, and hence $\sum_{n=1}^L w_{n,I,J}^\Delta \leq N\breve{W}^{\Delta,\epsilon}(B_{J,K}^\Delta(p^\Delta)) - (N-L)W_1^\Delta(B_{J,K}^\Delta(p^\Delta))$. So, for the preceding inequality to hold it is necessary that

$$
\begin{aligned}
&(1-\delta)L\bar{\alpha}^\Delta\left[\lambda(p^\Delta)h - s\right] + \delta\left\{\Pr^\Delta(\emptyset)Ls + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\Lambda_{0,K}^\Delta(p^\Delta)Ls\right. \\
&\qquad\left. + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\sum_{J=1}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)\left[N\breve{W}^{\Delta,\epsilon}(B_{J,K}^\Delta(p^\Delta)) - (N-L)W_1^\Delta(B_{J,K}^\Delta(p^\Delta))\right]\right\} \\
&\geq \delta\left\{\sum_{n=1}^{L}\Pr_{-n}^\Delta(\emptyset)s + \sum_{n=1}^{L}\sum_{K=1}^{L-1}\sum_{|I|=K,\, n\notin I}\Pr_{-n}^\Delta(I)\Lambda_{0,K}^\Delta(p^\Delta)s\right. \\
&\qquad\left. + \sum_{n=1}^{L}\sum_{K=1}^{L-1}\sum_{|I|=K,\, n\notin I}\Pr_{-n}^\Delta(I)\sum_{J=1}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)W_1^\Delta(B_{J,K}^\Delta(p^\Delta))\right\}.
\end{aligned}
$$

As

$$
\Pr^\Delta(\emptyset) + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I) = 1 \quad \text{and} \quad \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)K = L\bar{\alpha}^\Delta,
$$

we have the first-order expansions

$$
\begin{aligned}
&\Pr^\Delta(\emptyset) + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\Lambda_{0,K}^\Delta(p^\Delta) \\
&\quad = \Pr^\Delta(\emptyset) + \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\left(1 - K\lambda(p^\Delta)\Delta\right) + o(\Delta) \\
&\quad = 1 - L\bar{\alpha}^\Delta\lambda(p^\Delta)\Delta + o(\Delta)
\end{aligned}
$$

and

$$
\sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)\Lambda_{1,K}^\Delta(p^\Delta) = \sum_{K=1}^{L}\sum_{|I|=K}\Pr^\Delta(I)K\lambda(p^\Delta)\Delta + o(\Delta) = L\bar{\alpha}^\Delta\lambda(p^\Delta)\Delta + o(\Delta),
$$

so the left-hand side of the last inequality expands as

$$Ls + L\left\{r\bar{\alpha}\,[\lambda(\breve{p})h - s] - rs + \bar{\alpha}\lambda(\breve{p})\,[NV_{N,\breve{p}_\epsilon}(j(\breve{p})) - (N-L)V_1^*(j(\breve{p})) - Ls]\right\}\Delta + o(\Delta)$$

with $\bar{\alpha} = \lim_{\Delta\to 0}\bar{\alpha}^\Delta$. In the same way, the identities

$$\Pr{}_{-n}^{\Delta}(\emptyset) + \sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr{}_{-n}^{\Delta}(I) = 1 \quad \text{and} \quad \sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr{}_{-n}^{\Delta}(I)K = L\bar{\alpha}^\Delta - \alpha_n^\Delta$$

imply

$$\sum_{n=1}^{L}\Pr{}_{-n}^{\Delta}(\emptyset) + \sum_{n=1}^{L}\sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr{}_{-n}^{\Delta}(I)\Lambda_{0,K}^{\Delta}(p^\Delta) = L - L(L-1)\bar{\alpha}^\Delta\lambda(p^\Delta)\Delta + o(\Delta)$$

and

$$\sum_{n=1}^{L}\sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr{}_{-n}^{\Delta}(I)\Lambda_{1,K}^{\Delta}(p^\Delta) = L(L-1)\bar{\alpha}^\Delta\lambda(p^\Delta)\Delta + o(\Delta),$$

and so the right-hand side of the inequality expands as

$$Ls + L\left\{-rs + (L-1)\bar{\alpha}\lambda(\breve{p})\,[V_1^*(j(\breve{p})) - s]\right\}\Delta + o(\Delta).$$

Comparing terms of order Δ, dividing by L and letting $\epsilon \to 0$, we obtain

$$\bar{\alpha}\left\{\lambda(\breve{p})\,\big[NV_{N,\breve{p}}(j(\breve{p})) - (N-1)V_1^*(j(\breve{p})) - s\big] - rc(\breve{p})\right\} \geq 0.$$

By Lemma 1, this means $\breve{p} \geq \hat{p}$ whenever $\bar{\alpha} > 0$.

For the case that $\bar{\alpha} = 0$, we write the optimality condition for player $n \in \{1, \ldots, L\}$ as

$$\begin{aligned}
&(1-\delta)\lambda(p^\Delta)h + \delta\left\{\sum_{K=0}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr{}_{-n}^{\Delta}(I)\sum_{J=0}^{\infty}\Lambda_{J,K+1}^{\Delta}(p^\Delta)w_{n,I\dot\cup\{n\},J}^{\Delta}\right\}\\
\geq\ &(1-\delta)s + \delta\left\{\Pr{}_{-n}^{\Delta}(\emptyset)w_{n,\emptyset}^{\Delta} + \sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr{}_{-n}^{\Delta}(I)\sum_{J=0}^{\infty}\Lambda_{J,K}^{\Delta}(p^\Delta)w_{n,I,J}^{\Delta}\right\}.
\end{aligned}$$

As above, $w_{n,\emptyset}^\Delta \geq s$, and $w_{n,I,0}^\Delta = s$ whenever $I \neq \emptyset$. For $|I| = K > 0$ and $J > 0$, moreover, we have $w_{n,I,J}^\Delta \geq W_1^\Delta(B_{J,K}^\Delta(p^\Delta))$, $w_{n,I\dot\cup\{n\},J}^\Delta \geq W_1^\Delta(B_{J,K+1}^\Delta(p^\Delta))$ and $w_{n,I\dot\cup\{n\},J}^\Delta \leq N\breve{W}^{\Delta,\epsilon}(B_{J,K+1}^\Delta(p^\Delta)) - (N-1)W_1^\Delta(B_{J,K+1}^\Delta(p^\Delta))$. So, for the optimality condition to hold, it is necessary that

$$(1-\delta)\lambda(p^\Delta)h+\delta\left\{\sum_{K=0}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)\Lambda_{0,K+1}^\Delta(p^\Delta)s\right.$$

$$\left.+\sum_{K=0}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)\sum_{J=1}^{\infty}\Lambda_{J,K+1}^\Delta(p^\Delta)\left[N\breve{W}^{\Delta,\epsilon}(B_{J,K+1}^\Delta(p^\Delta))-(N-1)W_1^\Delta(B_{J,K+1}^\Delta(p^\Delta))\right]\right\}$$

$$\geq\ (1-\delta)s+\delta\left\{\Pr_{-n}^\Delta(\emptyset)s+\sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)\Lambda_{0,K}^\Delta(p^\Delta)s\right.$$

$$\left.+\sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)\sum_{J=1}^{\infty}\Lambda_{J,K}^\Delta(p^\Delta)W_1^\Delta(B_{J,K}^\Delta(p^\Delta))\right\}.$$

Now,

$$\sum_{K=1}^{L-1}\sum_{|I|=K,\,n\notin I}\Pr_{-n}^\Delta(I)K=L\bar{\alpha}^\Delta-\alpha_n^\Delta\to 0$$

as Δ vanishes. Therefore, the left-hand side of the above inequality expands as

$$s+\left\{r\left[\lambda(\breve{p})h-s\right]+\lambda(\breve{p})\left[NV_{N,\breve{p}_\epsilon}(j(\breve{p}))-(N-1)V_1^*(j(\breve{p}))-s\right]\right\}\Delta+o(\Delta),$$

and the right-hand side as $s+o(\Delta)$. Comparing terms of order Δ, letting $\epsilon\to 0$ and using Lemma 1 once more, we again obtain $\breve{p}\geq\hat{p}$.

Given that we have $\breve{p}=\hat{p}$, therefore, the proof is now easily completed along the lines of the proof of Proposition 5. ■

Proof of Lemma 3:

Simple algebra yields

$$\frac{j(p_N^*)}{p_1^*}=\frac{\lambda_1}{\lambda_0}\,\frac{\mu_N}{\mu_1}\,\frac{(\mu_1+1)(\lambda_1 h-s)+\mu_1(s-\lambda_0 h)}{(\mu_N+1)(\lambda_1 h-s)+(\lambda_1/\lambda_0)\mu_N(s-\lambda_0 h)}\,.$$

From the implicit definitions of μ_1 and μ_N, we obtain $\lim_{r\to 0}\mu_1=\lim_{r\to 0}\mu_N=0$ (so that the third fraction in the previous expression converges to 1) and

$$\lim_{r\to 0}\frac{\partial\mu_1}{\partial r}=\left[\lambda_1-\lambda_0+\lambda_0\ln\frac{\lambda_0}{\lambda_1}\right]^{-1}=N\lim_{r\to 0}\frac{\partial\mu_N}{\partial r}$$

implying

$$\lim_{r\to 0}\frac{\mu_N}{\mu_1}=\frac{1}{N}$$

by l'Hôpital's rule.

Furthermore, we note that we can write equivalently

$$\frac{j(p_N^*)}{p_1^*} = \frac{\lambda_1}{\lambda_0} \frac{(1+\frac{1}{\mu_1})(\lambda_1 h - s) + (s - \lambda_0 h)}{(1+\frac{1}{\mu_N})(\lambda_1 h - s) + (\lambda_1/\lambda_0)(s - \lambda_0 h)}\,.$$

As $\lim_{r\to\infty} \mu_1 = \lim_{r\to\infty} \mu_N = \infty$, we can immediately conclude that

$$\lim_{r\to\infty} \frac{j(p_N^*)}{p_1^*} = \frac{\lambda_1 h}{s}\,.$$

■

Proof of Lemma 4: For the case that $\hat{p} = p_N^*$, this is shown in Keller and Rady (2010). Thus, in what follows we shall assume that $\hat{p} > p_N^*$.

Recall the defining equation for $\hat{p}$ from Lemma 1,

$$\lambda(\hat{p}) N V_{N,\hat{p}}(j(\hat{p})) - \lambda(\hat{p}) s - rc(\hat{p}) = (N-1)\lambda(\hat{p}) V_1^*(j(\hat{p})).$$

We make use of the closed-form expression for $V_{N,\hat{p}}$ to rewrite its left-hand side as

$$N\lambda(\hat{p})\lambda(j(\hat{p}))h + Nc(\hat{p})[\lambda_0 - \mu_N(\lambda_1 - \lambda_0)] - \lambda(\hat{p})s.$$

Similarly, by noting that $\hat{p} > p_N^*$ implies $j(\hat{p}) > j(p_N^*) > p_1^*$, we can make use of the closed-form expression for V_1^* to rewrite the right-hand side as

$$(N-1)\lambda(\hat{p})\lambda(j(\hat{p}))h + (N-1)c(p_1^*)\frac{u(\hat{p};\mu_1)}{u(p_1^*;\mu_1)}[r + \lambda_0 - \mu_1(\lambda_1 - \lambda_0)].$$

Combining, we have

$$\frac{\lambda(\hat{p})\lambda(j(\hat{p}))h + Nc(\hat{p})[\lambda_0 - \mu_N(\lambda_1 - \lambda_0)] - \lambda(\hat{p})s}{(N-1)[r + \lambda_0 - \mu_1(\lambda_1 - \lambda_0)]c(p_1^*)} = \frac{u(\hat{p};\mu_1)}{u(p_1^*;\mu_1)}.$$

It is convenient to change variables to

$$\beta = \frac{\lambda_0}{\lambda_1} \quad \text{and} \quad y = \frac{\lambda_1}{\lambda_0}\frac{\lambda_1 h - s}{s - \lambda_0 h}\frac{\hat{p}}{1-\hat{p}}\,.$$

The implicit definitions of μ_1 and μ_N imply

$$N = \frac{\beta^{1+\mu_1} - \beta + \mu_1(1-\beta)}{\beta^{1+\mu_N} - \beta + \mu_N(1-\beta)}\,,$$

allowing us to rewrite the defining equation for $\hat{p}$ as the equation $F(y, \mu_N) = 0$ with

$$\begin{aligned} F(y,\mu) &= 1 - y + [\beta(1+\mu)y - \mu]\,\frac{1-\beta}{\beta}\,\frac{\beta^{1+\mu_1} - \beta + \mu_1(1-\beta)}{(\mu_1-\mu)(1-\beta) + \beta^{1+\mu_1} - \beta^{1+\mu}} \\ &\quad - \frac{\mu_1^{\mu_1}}{(1+\mu_1)^{1+\mu_1}}\, y^{-\mu_1}. \end{aligned}$$

As y is a strictly increasing function of $\hat{p}$, we know from Lemma 1 that $F(\cdot, \mu_N)$ admits a unique root, and that it is strictly increasing in a neighborhood of this root.

A straightforward computation shows that

$$\frac{\partial F(y,\mu_N)}{\partial \mu} = \frac{1-\beta}{\beta}\,\frac{\beta^{1+\mu_1} - \beta + \mu_1(1-\beta)}{((\mu_1-\mu_N)(1-\beta) + \beta^{1+\mu_1} - \beta^{1+\mu_N})^2}\,\zeta(y,\mu_N)$$

with

$$\zeta(y,\mu) = \beta(1-\beta)(1+\mu_1)y - (1-\beta)\mu_1 + (1-\beta y)(\beta^{1+\mu} - \beta^{1+\mu_1}) + \beta^{1+\mu}\,(\beta(1+\mu)y - \mu)\ln(\beta).$$

As $p_N^* < \hat{p} < p_1^*$, we have

$$\frac{\mu_N}{1+\mu_N} < \beta y < \frac{\mu_1}{1+\mu_1}\,,$$

which implies

$$\zeta(y,\mu_1) = (\beta(1+\mu_1)y - \mu_1)\,(1 - \beta + \beta^{1+\mu_1}\log(\beta)) < 0$$

and

$$\frac{\partial \zeta(y,\mu)}{\partial \mu} = \beta^{1+\mu}[\beta(1+\mu)y - \mu]\ln(\beta)^2 > 0$$

for all $\mu \in [\mu_N, \mu_1]$. This establishes $\zeta(y, \mu_N) < 0$.

By the implicit function theorem, therefore, y is increasing in μ_N. Recalling from Keller and Rady (2010) that μ_N is decreasing in N, we have thus shown that y (and hence $\hat{p}$) are decreasing in N. ■

B Convergence and Comparison Results

To establish uniform convergence of certain discrete-time value functions to their continuous-time limits, we will need the following result.[6]

Lemma B.1 *Let $\{T^\Delta\}_{\Delta>0}$ be a family of contraction mappings on the Banach space $(\mathcal{W}; \|\cdot\|)$ with moduli $\{\beta^\Delta\}_{\Delta>0}$ and associated fixed points $\{w^\Delta\}_{\Delta>0}$. Suppose that there is a constant $\rho > 0$ such that $1 - \beta^\Delta = \rho\Delta + o(\Delta)$ as $\Delta \to 0$. Then, a sufficient condition for w^Δ to converge in $(\mathcal{W}; \|\cdot\|)$ to the limit v as $\Delta \to 0$ is that $\|T^\Delta v - v\| = o(\Delta)$.*

[6]To the best of our knowledge, the earliest appearance of this result in the economics literature is in Biais et al. (2007). A related approach is taken in Sadzik and Stacchetti (2012).

PROOF: As

$$\|w^\Delta - v\| = \|T^\Delta w^\Delta - v\| \leq \|T^\Delta w^\Delta - T^\Delta v\| + \|T^\Delta v - v\| \leq \beta^\Delta \|w^\Delta - v\| + \|T^\Delta v - v\|,$$

the stated conditions on β^Δ and $\|T^\Delta v - v\|$ imply

$$\|w^\Delta - v\| \leq \frac{\|T^\Delta v - v\|}{1-\beta^\Delta} = \frac{\Delta f(\Delta)}{\rho\Delta + \Delta g(\Delta)} = \frac{f(\Delta)}{\rho + g(\Delta)}$$

with $\lim_{\Delta\to 0} f(\Delta) = \lim_{\Delta \to 0} g(\Delta) = 0$. ■

In our applications of this lemma, we shall take $\mathcal{W}$ to be the Banach space of bounded functions on the unit interval, equipped with the supremum norm. The operators T^Δ will be Bellman operators for certain optimal strategies in the experimentation game with period length Δ; the corresponding moduli will be $\beta^\Delta = \delta = e^{-r\Delta}$.

The limit functions will belong to the set $\mathcal{V}$ of all continuous $v \in \mathcal{W}$ with the following properties: there are finitely many beliefs $\{p_\ell\}_{\ell=0}^L$ with $0 = p_0 < p_1 < \ldots < p_{L-1} < p_L = 1$ such that for all $\ell = 1, \ldots, L$, (i) the function v is once continuously differentiable with bounded derivative v' on the interval $(p_{\ell-1}, p_\ell)$, (ii) $\lim_{p\uparrow p_\ell} v'(p)$ equals the left-hand derivative of v at p_ℓ, and (iii) $\lim_{p\downarrow p_{\ell-1}} v'(p)$ equals the right-hand derivative of v at $p_{\ell-1}$. In the following, we shall always take $v'(p_\ell)$ to mean the left-hand derivative at p_ℓ for $\ell \geq 1$, and the right-hand derivative for $\ell = 0$.

With this convention, the term

$$b(p,v) = \frac{\lambda(p)}{r}\,[v(j(p)) - v(p)] - \frac{\lambda_1 - \lambda_0}{r}\, p(1-p)\, v'(p)$$

is well-defined on the entire unit interval for any $v \in \mathcal{V}$. We can now provide a first-order expansion for the discounted expectation $\delta \mathcal{E}_K^\Delta$ that will appear in the Bellman operators of interest.[7]

Lemma B.2 *For $K \in \{0, 1, \ldots, N\}$ and $v \in \mathcal{V}$,*

$$\lim_{\Delta\to 0} \frac{1}{\Delta} \left\| \delta\, \mathcal{E}_K^\Delta v - v - r[K b(\cdot, v) - v]\Delta \right\| = 0.$$

PROOF: This follows from a straightforward Taylor expansion. ■

Our first application of Lemmas B.1 and B.2 concerns the upper bound on equilibrium payoffs introduced at the start of Section 5. Take $\tilde{p}$ as defined there. Given $\Delta > 0$, $\epsilon > 0$ and any bounded function w on $[0,1]$, define a bounded function $\widetilde{T}^{\Delta,\epsilon} w$ by

$$\widetilde{T}^{\Delta,\epsilon} w(p) = \begin{cases} \max\left\{(1-\delta)\lambda(p)h + \delta \mathcal{E}_N^\Delta w(p),\ (1-\delta)s + \delta w(p)\right\} & \text{if} \quad p > \tilde{p} - \epsilon, \\ (1-\delta)s + \delta w(p) & \text{if} \quad p \leq \tilde{p} - \epsilon. \end{cases}$$

[7] Up to discounting, this is nothing but the computation of the infinitesimal generator of the process of posterior beliefs, of course.

The operator $\widetilde{T}^{\Delta,\epsilon}$ satisfies Blackwell's sufficient conditions for being a contraction mapping with modulus δ on the Banach space $\mathcal{W}$ of bounded functions on $[0,1]$ equipped with the supremum norm $\|\cdot\|$: monotonicity ($v \leq w$ implies $\widetilde{T}^{\Delta,\epsilon}v \leq \widetilde{T}^{\Delta,\epsilon}w$) and discounting ($\widetilde{T}^{\Delta,\epsilon}(w+c) = \widetilde{T}^{\Delta,\epsilon}w + \delta c$ for any real number c). By the contraction mapping theorem, $\widetilde{T}^{\Delta,\epsilon}$ has a unique fixed point in $\mathcal{W}$; this is the value function $\widetilde{W}^{\Delta,\epsilon}$ of the constrained planner's problem considered in Section 5.

From Keller and Rady (2010), we know that the corresponding continuous-time value function is V_{N,p_ϵ} with $p_\epsilon = \max\{\tilde{p} - \epsilon, p_N^*\}$. It belongs to $\mathcal{V}$ and satisfies $V_{N,p_\epsilon}(p) = \lambda(p)h + Nb(p, V_{N,p_\epsilon}) > s$ on $(p_\epsilon, 1]$. For $p_\epsilon = p_N^*$, moreover, $\lambda(p)h + Nb(p, V_{N,p_\epsilon}) - s$ is zero at p_ϵ and negative on $[0, p_\epsilon)$.

Lemma B.3 $\widetilde{W}^{\Delta,\epsilon} \to V_{N,p_\epsilon}$ *uniformly as* $\Delta \to 0$.

Proof: To ease the notational burden, we write v instead of V_{N,p_ϵ}. Lemma B.2 then implies

$$\begin{aligned}(1-\delta)\lambda(p)h + \delta\mathcal{E}_N^\Delta v(p) &= v(p) + r\left[\lambda(p)h + Nb(p,v) - v(p)\right]\Delta + o(\Delta), \\ (1-\delta)s + \delta v(p) &= v(p) + r\left[s - v(p)\right]\Delta + o(\Delta).\end{aligned}$$

Suppose first that $p_\epsilon = \tilde{p} - \epsilon > p_N^*$. For $p > \tilde{p} - \epsilon$, we have $v(p) = \lambda(p)h + Nb(p,v) > s$, and hence $\widetilde{T}^{\Delta,\epsilon}v(p) = (1-\delta)\lambda(p)h + \delta\mathcal{E}_N^\Delta v(p) = v(p) + o(\Delta)$ for small Δ.

Next, suppose that $p_\epsilon = p_N^* \geq \tilde{p} - \epsilon$. For $p > p_N^*$, the same argument as in the previous paragraph yields $\widetilde{T}^{\Delta,\epsilon}v(p) = (1-\delta)\lambda(p)h + \delta\mathcal{E}_N^\Delta v(p) = v(p) + o(\Delta)$ for small Δ. For $p \in (\tilde{p} - \epsilon, p_N^*]$, we have $v(p) = s \geq \lambda(p)h + Nb(p,v)$, which once more implies $\widetilde{T}^{\Delta,\epsilon}v(p) = v(p) + o(\Delta)$ for small Δ.

As $\widetilde{T}^{\Delta,\epsilon}v(p) = s = v(p)$ trivially on $[0, \tilde{p} - \epsilon]$, we have established that $\|\widetilde{T}^{\Delta,\epsilon}v - v\| = o(\Delta)$. As the modulus of the contraction $\widetilde{T}^{\Delta,\epsilon}$ is $\delta = e^{-r\Delta} = 1 - r\Delta + o(\Delta)$, uniform convergence $\widetilde{W}^{\Delta,\epsilon} \to v$ now follows from Lemma B.1. ■

The second application of Lemmas B.1 and B.2 concerns the payoffs in the bad state of the equilibrium constructed in Section 6.3. Fix a cutoff $\bar{p} > p^m$, and let $K(p) = N - 1$ when $p > \bar{p}$, and $K(p) = 0$ otherwise. Given $\Delta > 0$, and any bounded function w on $[0,1]$, define a bounded function $\underline{T}^\Delta w$ by

$$\underline{T}^\Delta w(p) = \max\left\{(1-\delta)\lambda(p)h + \delta\mathcal{E}_{K(p)+1}^\Delta w(p),\ (1-\delta)s + \delta\mathcal{E}_{K(p)}^\Delta w(p)\right\}.$$

The operator $\underline{T}^\Delta$ also satisfies Blackwell's sufficient conditions for being a contraction mapping with modulus δ on $\mathcal{W}$. Its unique fixed point in this space is the payoff function $\underline{w}^\Delta$ (introduced in Section 6.3) from playing a best response against $N-1$ opponents who all play risky when $p > \bar{p}$, and safe otherwise. For $\bar{p} = 1$, the fixed point is the single-agent value function W_1^Δ.

In Section 6.3, we introduced the notation $V_{1,\bar{p}}$ for the continuous-time counterpart to this payoff function. The methods employed in Keller and Rady (2010) can be used to establish that $V_{1,\bar{p}}$ has the following properties. First, there is a cutoff $p^\dagger < p^m$ such that $V_{1,\bar{p}} = s$ on $[0, p^\dagger]$, and $V_{1,\bar{p}} > s$ everywhere else. Second, $V_{1,\bar{p}} \in \mathcal{V}$, being continuously differentiable everywhere except at $\bar{p}$. Third, $V_{1,\bar{p}}$ solves the Bellman equation

$$v(p) = \max\left\{\lambda(p)h + [K(p)+1]b(p,v),\ s + K(p)b(p,v)\right\}.$$

Fourth, because of smooth pasting at $p^\dagger$, the term $\lambda(p)h + b(p, V_{1,\bar{p}}) - s$ is continuous in p except at $\bar{p}$; it has a single zero at $p^\dagger$, being positive to the right of it and negative to the left. Finally, we note that $V_{1,\bar{p}} = V_1^*$ and $p^\dagger = p_1^*$ for $\bar{p} = 1$.

Let $p^{\dagger,\Delta} = \inf\{p \colon \underline{w}^\Delta(p) > s\}$.

Lemma B.4 *$\underline{w}^\Delta \to V_{1,\bar{p}}$ uniformly as $\Delta \to 0$, and $\liminf_{\Delta\to 0} p^{\dagger,\Delta} = p^\dagger$.*

PROOF: To ease the notational burden, we write v instead of $V_{1,\bar{p}}$.

For $p > \bar{p}$, we have $K(p) = N-1$, and Lemma B.2 implies

$$\begin{aligned}(1-\delta)\lambda(p)h + \delta\mathcal{E}^\Delta_{K(p)+1}v(p) &= v(p) + r\left[\lambda(p)h + Nb(p,v) - v(p)\right]\Delta + o(\Delta),\\ (1-\delta)s + \delta\mathcal{E}^\Delta_{K(p)}v(p) &= v(p) + r\left[s + (N-1)b(p,v) - v(p)\right]\Delta + o(\Delta).\end{aligned}$$

As $v(p) = \lambda(p)h + Nb(p,v) > s + (N-1)b(p,v)$, we thus have $\underline{T}^\Delta v(p) = (1-\delta)\lambda(p)h + \delta\mathcal{E}^\Delta_{K(p)+1}v(p) = v(p) + o(\Delta)$ for small Δ.

On $(p^\dagger, \bar{p}]$, we have $K(p) = 0$ and

$$\begin{aligned}(1-\delta)\lambda(p)h + \delta\mathcal{E}^\Delta_{K(p)+1}v(p) &= v(p) + r\left[\lambda(p)h + b(p,v) - v(p)\right]\Delta + o(\Delta),\\ (1-\delta)s + \delta\mathcal{E}^\Delta_{K(p)}v(p) &= v(p) + r\left[s - v(p)\right]\Delta + o(\Delta).\end{aligned}$$

As $v(p) = \lambda(p)h + b(p,v) > s$, we again have $\underline{T}^\Delta v(p) = (1-\delta)\lambda(p)h + \delta\mathcal{E}^\Delta_{K(p)+1}v(p) = v(p) + o(\Delta)$ for small Δ.

For $p \le p^\dagger$, finally, we have $K(p) = 0$ and $v(p) = s$, hence

$$\begin{aligned}(1-\delta)\lambda(p)h + \delta\mathcal{E}^\Delta_{K(p)+1}v(p) &= s + r\left[\lambda(p)h + b(p,v) - v(p)\right]\Delta + o(\Delta),\\ (1-\delta)s + \delta\mathcal{E}^\Delta_{K(p)}v(p) &= s.\end{aligned}$$

As $v(p) = s \ge \lambda(p)h + b(p,v)$, this once more implies $\underline{T}^\Delta v(p) = v(p) + o(\Delta)$ for small Δ.

We have thus shown that $\|\underline{T}^\Delta v - v\| = o(\Delta)$. Uniform convergence $\underline{w}^\Delta \to v$ now follows from Lemma B.1.

Turning to the second part of the lemma, we define $p^{\dagger,0} = \liminf_{\Delta\to 0} p^{\dagger,\Delta}$. For a sequence of Δ's converging to 0 such that the corresponding beliefs $p^{\dagger,\Delta}$ converge to $p^{\dagger,0}$, choose $p^\Delta > p^{\dagger,\Delta}$ such that $B^\Delta_{0,1}(p^\Delta) < p^{\dagger,\Delta}$. Along the sequence, we have $\underline{w}^\Delta(p^\Delta) > s = \underline{w}^\Delta(B^\Delta_{0,1}(p^\Delta))$

and $(1-\delta)\lambda(p^\Delta)h + \delta\mathcal{E}_1^\Delta \underline{w}^\Delta(p^\Delta) > (1-\delta)s + \delta\underline{w}^\Delta(p^\Delta) > s$. As

$$\begin{aligned}
&(1-\delta)\lambda(p^\Delta)h + \delta\mathcal{E}_1^\Delta \underline{w}^\Delta(p^\Delta) \\
&= r\Delta\,\lambda(p^\Delta)h + (1-r\Delta)\left\{(1-\lambda(p^\Delta)\Delta)s + \lambda(p^\Delta)\Delta\,\underline{w}^\Delta\big(B_{1,1}^\Delta(p^\Delta)\big)\right\} + o(\Delta) \\
&= s + \left\{r[\lambda(p^{\dagger,0})h - s] + \lambda(p^{\dagger,0})[v(j(p^{\dagger,0})) - s]\right\}\Delta + o(\Delta),
\end{aligned}$$

we can conclude that $\lambda(p^{\dagger,0})[v(j(p^{\dagger,0})) - s] \geq rc(p^{\dagger,0})$. As $v'(p) = 0$ and $\lambda(p)[v(j(p)) - s] = rb(p,v) < rc(p)$ for $p < p^\dagger$, this implies $p^{\dagger,0} \geq p^\dagger$. And since the inequality $p^{\dagger,0} > p^\dagger$ would imply $v(p) > s = \lim_{\Delta\to 0}\underline{w}^\Delta(p)$ immediately to the right of $p^\dagger$, we must have $p^{\dagger,0} = p^\dagger$. ■

Our third uniform convergence result also concerns the continuous-time limits of equilibrium payoffs in the bad state. As it is straightforward to establish with the methods used in Keller and Rady (2010), we state it without proof.

Lemma B.5 $V_{1,\bar{p}} \to V_1^*$ *uniformly as* $\bar{p} \to 1$. *The convergence is monotone in the sense that* $\bar{p}' > \bar{p}$ *implies* $V_{1,\bar{p}'} < V_{1,\bar{p}}$ *on* $\{p\colon s < V_{1,\bar{p}}(p) < \lambda_1 h\}$.

Our last result on uniform convergence concerns the payoffs in the good state of the equilibrium constructed in Section 6.3. Fix a cutoff $\underline{p}$ and consider the strategy profile where all N players play risky for $p > \underline{p}$, and all play safe otherwise. As in Section 6.3, we write $\overline{w}^\Delta$ for the players' common payoff function from this strategy profile when actions are frozen for a length of time Δ. The corresponding payoff function in continuous time is $V_{N,\underline{p}}$. The following result can be obtained from first principles; its proof does not rely on Lemmas B.1 and B.2.

Lemma B.6 $\overline{w}^\Delta \to V_{N,\underline{p}}$ *uniformly as* $\Delta \to 0$.

PROOF: As $\overline{w}^\Delta(p) = V_{N,\underline{p}}(p) = s$ for $p \leq \underline{p}$, there is nothing to show for these beliefs. Fix an initial belief $p > \underline{p}$, therefore, and consider the process of beliefs $\{p_t\}$ starting from $p_0 = p$ that corresponds to N players using the risky arm. Let $\tau = \inf\{t \geq 0\colon p_t \leq \underline{p}\}$ and $\tau^\Delta = \inf\{t = \Delta, 2\Delta, 3\Delta, \ldots\colon p_t \leq \underline{p}\}$. Then,

$$\begin{aligned}
V_{N,\underline{p}}(p) &= \mathbb{E}\left[\int_0^\tau re^{-rt}h\,dN_{\theta,t} + e^{-r\tau}s\right], \\
\overline{w}^\Delta(p) &= \mathbb{E}\left[\int_0^{\tau^\Delta} re^{-rt}h\,dN_{\theta,t} + e^{-r\tau^\Delta}s\right]
\end{aligned}$$

where N_θ is a Poisson process with intensity λ_θ. As $\tau \leq \tau^\Delta \leq \tau + \Delta$ almost surely, we have

$$\begin{aligned}
|\overline{w}^\Delta(p) - V_{N,\underline{p}}(p)| &\leq \mathbb{E}\left[\int_\tau^{\tau^\Delta} re^{-rt}h\,dN_{\theta,t} + |e^{-r\tau^\Delta} - e^{-r\tau}|s\right] \\
&\leq \mathbb{E}\left[\int_0^\Delta re^{-rt}h\,dN_{1,t}\right] + (1-e^{-r\Delta})s \\
&= (1-e^{-r\Delta})(\lambda_1 h + s),
\end{aligned}$$

and hence $\lim_{\Delta\to 0}\|\overline{w}^\Delta - V_{N,\underline{p}}\| = 0$ as claimed. ■

The remaining auxiliary results needed for the proof of Proposition 3 are comparison results for $\overline{w}^\Delta$ and $\underline{w}^\Delta$ as Δ becomes small. We start with equilibrium payoffs in the good state.

Lemma B.7 *Let $\underline{p} > p_N^*$. Then $\overline{w}^\Delta \geq V_{N,\underline{p}}$ for Δ sufficiently small.*

PROOF: In addition to the stopping times τ and τ^Δ introduced in the proof of Lemma B.6, we define $\tau^* = \inf\{t \geq 0\colon p_t \leq p_N^*\}$, which is the stopping time that an N-player cooperative would use in continuous time. As $\underline{p} > p_N^*$, we have $\tau^* \geq \tau + \Delta^*$ where $\Delta^* > 0$ is the (deterministic) length of time needed for the posterior belief to decay from $\underline{p}$ to p_N^* when no lump sum arrives. For $\Delta \leq \Delta^*$, therefore, we have $\tau \leq \tau^\Delta \leq \tau + \Delta \leq \tau^*$, so τ^Δ yields an expected payoff no smaller than τ; that is, $\overline{w}^\Delta \geq V_{N,\underline{p}}$. ■

Turning to equilibrium payoffs in the bad state, we define

$$p^\flat = \frac{\mu^\flat(s-\lambda_0 h)}{(\mu^\flat+1)(\lambda_1 h - s) + \mu^\flat(s-\lambda_0 h)},$$

where

$$\mu^\flat = \mu_N + \frac{(N-1)r}{N(\lambda_1-\lambda_0)}\,.$$

Lemma B.8 *For $\underline{p} < p^\flat$ and Δ sufficiently small, $\underline{w}^\Delta \leq V_{N,\underline{p}}$.*

PROOF: To ease the notational burden, we write v instead of $V_{N,\underline{p}}$. It suffices to show that $\underline{T}^\Delta v \leq v$ for sufficiently small Δ.

Recall that for $p > \underline{p}$, $v(p) = \lambda(p)h + Cu(p)$ with $u(p) = (1-p)\left(\frac{1-p}{p}\right)^{\mu_N}$ where the constant $C > 0$ is chosen to ensure continuity at $\underline{p}$. It follows from Keller and Rady (2010) that v is strictly increasing on $[\underline{p}, 1]$.

The function u is strictly decreasing and strictly convex, and a straightforward computation reveals that $\delta\mathcal{E}_K^\Delta u(p) = \delta^{1-\frac{K}{N}}u(p)$ for all $\Delta > 0$, $K \in \{1,\ldots,N\}$ and $p \in (0,1]$. We further note that $\mathcal{E}_K^\Delta \lambda(p) = \lambda(p)$ for all K by the martingale property of beliefs.

We define a belief $\check{p}^\Delta$ by requiring that $B_{0,1}^\Delta(\check{p}^\Delta) = \underline{p}$. Starting from $p > \check{p}^\Delta$, when one player experiments for a length of time Δ without receiving a lump sum, the resulting posterior belief remains above $\underline{p}$.

On $(\bar{p}, 1]$, we now have

$$\begin{aligned}
\underline{T}^\Delta v(p) &= \max\left\{(1-\delta)\lambda(p)h + \delta\mathcal{E}_N^\Delta v(p),\ (1-\delta)s + \delta\mathcal{E}_{N-1}^\Delta v(p)\right\} \\
&= \max\left\{(1-\delta)\lambda(p)h + \delta\lambda(p)h + C\delta\mathcal{E}_N^\Delta u(p),\ (1-\delta)s + \delta\lambda(p)h + C\delta\mathcal{E}_{N-1}^\Delta u(p)\right\} \\
&= \lambda(p)h + C\delta\mathcal{E}_N^\Delta u(p) \\
&= v(p).
\end{aligned}$$

The third equality holds because $\delta\mathcal{E}_N^\Delta u(p) > \delta\mathcal{E}_{N-1}^\Delta u(p)$ (by strict convexity of u) and $\lambda(p)h > s$ (as $\bar{p} > p^m$ by assumption), the fourth holds because $\delta\mathcal{E}_N^\Delta u(p) = u(p)$.

On $(\check{p}^\Delta, \bar{p}]$, we have

$$
\begin{aligned}
\underline{T}^\Delta v(p) &= \max\Big\{(1-\delta)\lambda(p)h + \delta\mathcal{E}_1^\Delta v(p),\ (1-\delta)s + \delta v(p)\Big\} \\
&= \max\Big\{\lambda(p)h + C\delta\mathcal{E}_1^\Delta u(p),\ (1-\delta)s + \delta v(p)\Big\} \\
&< v(p),
\end{aligned}
$$

with the inequality holding because $\delta\mathcal{E}_1^\Delta u(p) = \delta^{\frac{N-1}{N}} u(p) < u(p)$ and $s < v(p)$.

On $(\underline{p}, \check{p}^\Delta]$, we still have $(1-\delta)s + \delta v(p) < v(p)$, while

$$
\begin{aligned}
&(1-\delta)\lambda(p)h + \delta\mathcal{E}_1^\Delta v(p) \\
&= (1-\delta)\lambda(p)h + \delta\Lambda_{0,1}^\Delta(p)\, s + \delta\sum_{j=1}^{\infty}\Lambda_{j,1}^\Delta(p)\, v(B_{j,1}^\Delta(p)) \\
&= (1-\delta)\lambda(p)h + \delta\Lambda_{0,1}^\Delta(p)\left[s - \lambda(B_{0,1}^\Delta(p))h - Cu(B_{0,1}^\Delta(p))\right] + \delta\mathcal{E}_1^\Delta[\lambda h + Cu](p) \\
&= \lambda(p)h + \delta\Lambda_{0,1}^\Delta(p)\left[s - \lambda(B_{0,1}^\Delta(p))h - Cu(B_{0,1}^\Delta(p))\right] + C\delta^{1-\frac{1}{N}}u(p) \\
&= v(p) + \delta F(p,\Delta)
\end{aligned}
$$

with

$$
F(p,\Delta) = C(\delta^{-\frac{1}{N}} - \delta^{-1})u(p) + \Lambda_{0,1}^\Delta(p)\left[s - \lambda(B_{0,1}^\Delta(p))h - Cu(B_{0,1}^\Delta(p))\right].
$$

As $\delta^{-\frac{1}{N}} = e^{r\Delta/N} < e^{r\Delta} = \delta^{-1}$, we have $F(\check{p}^\Delta, \Delta) < 0$. Moreover, as $\Lambda_{0,1}^\Delta(p) = p\gamma_1 + (1-p)\gamma_0$ and $B_{0,1}^\Delta(p) = p\gamma_1/\Lambda_{0,1}^\Delta(p)$, we have

$$
\Lambda_{0,1}^\Delta(p)\,\lambda(B_{0,1}^\Delta(p)) = p\lambda_1\gamma_1 + (1-p)\lambda_0\gamma_0
$$

and

$$
\Lambda_{0,1}^\Delta(p)\,u(B_{0,1}^\Delta(p)) = \gamma_0\left(\frac{\gamma_0}{\gamma_1}\right)^{\mu_N} u(p),
$$

hence

$$
F(p,\Delta) = C\left[\delta^{-\frac{1}{N}} - \delta^{-1} - \gamma_0\left(\frac{\gamma_0}{\gamma_1}\right)^{\mu_N}\right]u(p) + [p\gamma_1 + (1-p)\gamma_0]s - [p\lambda_1\gamma_1 + (1-p)\lambda_0\gamma_0]h,
$$

which is continuously differentiable at any $(p,\Delta) \in (0,1) \times I\!R$. For $\Delta \geq 0$, the nonlinear part of F is a negative multiple of u, so F is strictly concave in p. As $F_p(\underline{p}, 0) = -Cu'(\underline{p}) - \lambda'(\underline{p})h = -v'(\underline{p}+) < 0$, we see that for sufficiently small $\Delta > 0$, $F_p(\underline{p}, \Delta) < 0$ and hence $F(p,\Delta) < F(\underline{p},\Delta)$ for $p > \underline{p}$. As $F(\underline{p}, 0) = -Cu(\underline{p}) + s - \lambda(\underline{p})h = s - v(\underline{p}) = 0$, we thus have $\underline{T}^\Delta v < v$ on $(\underline{p}, \check{p}^\Delta]$ for sufficiently small Δ if we can show that $F_\Delta(\underline{p}, 0) < 0$. Computing

$$
F_\Delta(\underline{p}, 0) = \left[\tfrac{r}{N} - r + \lambda_0 - \mu_N(\lambda_1 - \lambda_0)\right](s - \lambda(\underline{p})h) + (\underline{p}\lambda_1^2 + (1-\underline{p})\lambda_0^2)h - \lambda(\underline{p})s,
$$

it is straightforward to check that $F_\Delta(\underline{p},0) < 0$ if and only if $\underline{p} < p^\flat$.

On $[0,\underline{p}]$, finally, the monotonicity of v on $[\underline{p},1]$ implies that $\mathcal{E}_1^\Delta v(p)$ is increasing in p. We thus have

$$(1-\delta)\lambda(p)h + \delta\mathcal{E}_1^\Delta v(p) \le (1-\delta)\lambda(\underline{p})h + \delta\mathcal{E}_1^\Delta v(\underline{p}) = v(\underline{p}) + \delta F(\underline{p},\Delta) < v(\underline{p}) = s$$

and hence $\underline{T}^\Delta v(p) = s = v(p)$. ■

Lemma B.9 *If $\lambda_0 > 0$, then $\hat{p} < p^\flat < p_1^*$.*

Proof: As $\mu_N < \mu_1$ and

$$r + \lambda_0 - \mu^\flat(\lambda_1 - \lambda_0) = \frac{r}{N} + \lambda_0 - \mu_N(\lambda_1-\lambda_0) = \lambda_0\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu_N} > \lambda_0\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu_1},$$

we have $\mu^\flat < \mu_1$. Combined with the fact that $\mu^\flat > \mu_N$, this implies $p_N^* < p^\flat < p_1^*$, which is already the desired result in the case that $j(p_N^*) \le p_1^*$ and $\hat{p} = p_N^*$.

Suppose therefore that $j(p_N^*) > p_1^*$ and $\hat{p} > p_N^*$. From Lemma 1, we know that $p^\flat > \hat{p}$ if and only if

$$\lambda(p^\flat)[NV_{N,p^\flat}(j(p^\flat)) - (N-1)V_1^*(j(p^\flat)) - s] - rc(p^\flat) > 0.$$

Arguing as in the proof of that lemma, we can rewrite the left-hand side of this inequality as

$$[p^\flat\lambda_1^2 + (1-p^\flat)\lambda_0^2]h + N\lambda_0\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu_N} c(p^\flat) - (N-1)\lambda_0\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu_1}\frac{c(p_1^*)}{u(p_1^*;\mu_1)}u(p^\flat;\mu_1) - \lambda(p^\flat)s - rc(p^\flat).$$

From the proof of Lemma B.8, moreover, we know that $F_\Delta(p^\flat,0) = 0$, which is equivalent to

$$[p^\flat\lambda_1^2 + (1-p^\flat)\lambda_0^2]h + \lambda_0\left(\frac{\lambda_0}{\lambda_1}\right)^{\mu_N} c(p^\flat) - \lambda(p^\flat)s - rc(p^\flat) = 0.$$

Thus, $p^\flat > \hat{p}$ if and only if

$$\frac{[r + \lambda_0 - \mu^\flat(\lambda_1-\lambda_0)]\,c(p^\flat)}{u(p^\flat;\mu_1)} > \frac{[r+\lambda_0 - \mu_1(\lambda_1-\lambda_0)]\,c(p_1^*)}{u(p_1^*;\mu_1)}.$$

Now, for $\mu > 0$ and

$$p(\mu) = \frac{\mu(s-\lambda_0 h)}{(\mu+1)(\lambda_1 h - s) + \mu(s-\lambda_0 h)},$$

a straightforward computation reveals that

$$\frac{c(p(\mu))}{u(p(\mu);\mu_1)} = \frac{(s-\lambda_0 h)\left(\frac{s-\lambda_0 h}{\lambda_1 h - s}\right)^{\mu_1}}{(\mu+1)\left(\frac{\mu+1}{\mu}\right)^{\mu_1}}.$$

Applying this to $p^\flat = p(\mu^\flat)$ and $p_1^* = p(\mu_1)$, we see that $p^\flat > \hat{p}$ if and only if the function

$$g(\mu) = \frac{r + \lambda_0 - \mu(\lambda_1 - \lambda_0)}{(\mu + 1)\left(\frac{\mu+1}{\mu}\right)^{\mu_1}}$$

satisfies $g(\mu^\flat) > g(\mu_1)$.

It is straightforward to show that $g'(\mu)$ has the same sign as $\mu^* - \mu$ where

$$\mu^* = \frac{\mu_1(r + \lambda_0)}{r + \lambda_1 + \mu_1(\lambda_1 - \lambda_0)} < \mu_1.$$

It is thus enough to show that $\mu^\flat > \mu^*$. Our assumption that $j(p_N^*) > p_1^*$ translates into

$$\mu_N > \frac{\mu_1 \lambda_0}{\lambda_1 + \mu_1(\lambda_1 - \lambda_0)}.$$

As $\frac{N-1}{N} \geq \frac{1}{2}$, this implies that $\mu^\flat$ is greater than

$$\bar{\mu} = \frac{\mu_1 \lambda_0}{\lambda_1 + \mu_1(\lambda_1 - \lambda_0)} + \frac{r}{2(\lambda_1 - \lambda_0)}.$$

The proof is complete, therefore, if we can show that $\bar{\mu} > \mu^*$.

Simple algebra shows that this inequality is equivalent to the concave quadratic

$$q(\mu) = \lambda_1(r + \lambda_1) + (\lambda_1 - \lambda_0)(r + 2\lambda_0)\mu - (\lambda_1 - \lambda_0)^2\mu^2$$

being positive at μ_1. We know from Keller and Rady (2010) that $\frac{r}{\lambda_1 - \lambda_0} < \mu_1 < \frac{r+\lambda_0}{\lambda_1 - \lambda_0}$. As $q(\frac{r}{\lambda_1 - \lambda_0}) = \lambda_1(r + \lambda_1) + 2\lambda_0 r$ and $q(\frac{r+\lambda_0}{\lambda_1 - \lambda_0}) = \lambda_1(r + \lambda_1) + \lambda_0(r + \lambda_0)$ are both positive, we can indeed conclude that $q(\mu_1) > 0$. ■

C Proofs for the Fully Revealing Case ($\lambda_0 = 0$)

Modifying notation slightly, we write Λ for the probability that, conditional on $\theta = 1$, a player has at least one success on his own risky arm in any given round, and g for the corresponding expected payoff per unit of time.[8]

Consider an SSE played at a given prior p, with associated payoff W. If $K \geq 1$ players unsuccessfully choose the risky arm, the belief jumps down to a posterior denoted p_K. Note that an SSE allows the continuation play to depend on the identity of these players. Taking the expectation over all possible combinations of K players who experiment, however, we can associate with each posterior p_K, $K \geq 1$, an expected continuation payoff W_K. If $K = 0$, so that no player experiments, the belief does not evolve, but there is no reason that the continuation strategies (and so the payoff) should remain the same. We denote

[8] *I.e.*, $\Lambda = 1 - e^{-\lambda_1 \Delta} = 1 - \gamma_1$ and $g = \lambda_1 h$.

the corresponding payoff by W_0. In addition, we write $\alpha \in [0,1]$ for the probability with which each player experiments at p, and Q_K for the probability that at least one player has a success, given p, when K of them experiment. The players' common payoff must then satisfy the following optimality equation:

$$\begin{aligned} W &= \max\Bigg\{(1-\delta)p_0 g + \delta \sum_{K=0}^{N-1} \binom{N-1}{K} \alpha^K (1-\alpha)^{N-1-K} [Q_{K+1} g + (1-Q_{K+1}) W_{K+1})]\ , \\ &\quad (1-\delta)s + \delta \sum_{K=1}^{N-1} \binom{N-1}{K} \alpha^K (1-\alpha)^{N-1-K} (Q_K g + (1-Q_K) W_K) + \delta (1-\alpha)^{N-1} W_0)\Bigg\}. \end{aligned}$$

The first term corresponds to the payoff from playing risky, the second from playing safe.

As it turns out, it is more convenient to work with odds ratios

$$l = \frac{p}{1-p} \quad \text{and} \quad l_K = \frac{p_K}{1-p_K}$$

which we refer to as "belief" as well. Note that

$$p_K = \frac{p\,(1-\Lambda)^K}{p\,(1-\Lambda)^K + 1 - p}$$

implies that $l_K = (1-\Lambda)^K l$. Note also that

$$1 - Q_K = p\,(1-\Lambda)^K + 1 - p = (1-p)(1+l_K), \quad Q_K = p - (1-p) l_K = (1-p)(l - l_K).$$

We define

$$m = \frac{s}{g-s}, \quad \omega = \frac{W-s}{(1-p)(g-s)}, \quad \omega_K = \frac{W_K - s}{(1-p_K)(g-s)}.$$

Note that $\omega \geq 0$ in any equilibrium, as s is a lower bound on the value. Simple computations now give

$$\begin{aligned} \omega &= \max\Bigg\{l - (1-\delta)m + \delta \sum_{K=0}^{N-1} \binom{N-1}{K} \alpha^K (1-\alpha)^{N-1-K} (\omega_{K+1} - l_{K+1})\ , \\ &\quad \delta l + \delta \sum_{K=0}^{N-1} \binom{N-1}{K} \alpha^K (1-\alpha)^{N-1-K} (\omega_K - l_K)\Bigg\}. \end{aligned}$$

It is also useful to introduce $w = \omega - l$ and $w_K = \omega_K - l_K$. We then get

$$\begin{aligned} w &= \max\Bigg\{-(1-\delta)m + \delta \sum_{K=0}^{N-1} \binom{N-1}{K} \alpha^K (1-\alpha)^{N-1-K} w_{K+1}\ , \\ &\quad -(1-\delta)l + \delta \sum_{K=0}^{N-1} \binom{N-1}{K} \alpha^K (1-\alpha)^{N-1-K} w_K\Bigg\}. \end{aligned} \tag{C.7}$$

We define

$$l^* = \frac{m}{1 + \frac{\delta}{1-\delta}\Lambda}.$$

This is the odds ratio corresponding to the single-agent cutoff p_1^Δ, *i.e.*, $l^* = p_1^\Delta/(1-p_1^\Delta)$. Note that $p_1^\Delta > p_1^*$ for $\Delta > 0$.

We are now ready to prove Lemma 2, which establishes that no perfect Bayesian equilibrium involves experimentation below p_1^Δ or, in terms of odds ratios, l^*.

Proof of Lemma 2: Let $\underline{l}$ be the infimum over all beliefs for which a positive probability of experimentation by some player can be implemented in a perfect Bayesian equilibrium. Note that $\underline{l} > 0$: This is because the social planner's solution is a cutoff policy, with cutoff bounded away from 0. Below this cutoff, s is both the minmax payoff of a player (which he can secure by always playing safe) and the highest average payoff that is feasible (given that this is the social optimum). Hence this must be the unique perfect Bayesian equilibrium payoff, and the unique policy that achieves it (from the social planner's problem) specifies that all players play safe.

Consider some prior belief $l \in [\underline{l}, \underline{l}/(1-\Lambda))$, so that a single failed experiment takes the posterior belief below $\underline{l}$, and fix an equilibrium in which at least one player experiments with positive probability in the first period. Let this be player n. As the normalized equilibrium payoff w at the belief l is bounded below by $-l$, and since by construction the payoff equals $-l_K$ at any belief l_K for $K \geq 1$, player n's payoff from playing safe is at least

$$-(1-\delta)l - \delta \sum_{I \subset N\setminus\{n\}} \prod_{i\in I}\alpha_i \prod_{i\in N\setminus(I\cup\{n\})} (1-\alpha_i)\; l_{|I|},$$

while the payoff from playing risky is

$$-(1-\delta)m - \delta \sum_{I \subset N\setminus\{n\}} \prod_{i\in I}\alpha_i \prod_{i\in N\setminus(I\cup\{n\})} (1-\alpha_i)\; l_{|I|+1}.$$

Thus, we must have

$$\begin{aligned}(1-\delta)(m-l) &\leq \delta \sum_{I \subset N\setminus\{n\}} \prod_{i\in I}\alpha_i \prod_{i\in N\setminus(I\cup\{n\})} (1-\alpha_i)\; (l_{|I|} - l_{|I|+1}) \\ &= \delta\Lambda l \sum_{I \subset N\setminus\{n\}} (1-\Lambda)^{|I|} \prod_{i\in I}\alpha_i \prod_{i\in N\setminus(I\cup\{n\})} (1-\alpha_i) \\ &\leq \delta\Lambda l.\end{aligned}$$

(The sum in the second line achieves its maximum of 1 when $\alpha_i = 0$ for all $i \neq n$.) This implies

$$l \geq \frac{m}{1 + \frac{\delta}{1-\delta}\Lambda} = l^*$$

and hence $\underline{l} \geq l^*$, establishing the lemma. ■

For all beliefs $l < l^*$, therefore, any equilibrium has $w = -l$, or $\omega = 0$, for each player. We now turn to the proof of Proposition 4.

PROOF OF PROPOSITION 4: Following terminology from repeated games, we say that we can *enforce* action $\alpha \in \{0, 1\}$ at belief l if we can construct an SSE for the prior belief l in which players prefer to choose α in the first round rather than deviate unilaterally.

Our first step is to derive sufficient conditions for enforcement of $\alpha \in \{0, 1\}$. The conditions to enforce these actions are intertwined, and must be derived simultaneously.

To enforce $\alpha = 0$ at l, it suffices that one round of using the safe arm followed by the best equilibrium payoff at l exceeds the payoff from one round of using the risky arm followed by the resulting continuation payoff at belief l_1 (as only the deviating player will have experimented). See below for the precise condition.

What does it take to enforce $\alpha = 1$ at l? If a player deviates to $\alpha = 0$, we jump to w_{N-1} rather than w_N in case all experiments fail. Assume that at l_{N-1} we can enforce $\alpha = 0$. As explained above, this implies that at l_{N-1}, a player's continuation payoff can be pushed down to what he would get by unilaterally deviating to experimentation, which is at most $-(1-\delta)m+\delta w_N$ where w_N is the highest possible continuation payoff at belief l_N. To enforce $\alpha = 1$ at l, it then suffices that

$$w = -(1-\delta)m + \delta w_N \geq -(1-\delta)l + \delta(-(1-\delta)m + \delta w_N),$$

with the same continuation payoff w_N on the left-hand side of the inequality. The inequality simplifies to

$$\delta w_N \geq (1-\delta)m - l;$$

by the formula for w, this is equivalent to $w \geq -l$, *i.e.*, $\omega \geq 0$. Given that

$$\omega = l - (1-\delta)m + \delta(\omega_N - l_N) = (1 - \delta(1-\Lambda)^N)l - (1-\delta)m + \delta\omega_N,$$

to show that $\omega \geq 0$, it thus suffices that

$$l \geq \frac{m}{1 + \frac{\delta}{1-\delta}(1-(1-\Lambda)^N)} = \tilde{l},$$

and that $\omega_N \geq 0$, which is necessarily the case if ω_N is an equilibrium payoff. Note that $(1-\Lambda)^N \tilde{l} \leq l^*$, so that $l_N \geq l^*$ implies $l \geq \tilde{l}$. In summary, to enforce $\alpha = 1$ at l, it suffices that $l_N \geq l^*$ and $\alpha = 0$ be enforceable at l_{N-1}.

How about enforcing $\alpha = 0$ at l? Suppose we can enforce it at $l_1, l_2, \ldots, l_{N-1}$, and that $l_N \geq l^*$. Note that $\alpha = 1$ is then enforceable at l from our previous argument, given our hypothesis that $\alpha = 0$ is enforceable at l_{N-1}. It then suffices that

$$-(1-\delta)l + \delta(-(1-\delta)m + \delta w_N) \geq -(1-\delta^N)m + \delta^N w_N,$$

where again it suffices that this holds for the highest value of w_N. To understand this expression, consider a player who deviates by experimenting. Then the following period the belief is down one step, and if $\alpha = 0$ is enforceable at l_1, it means that his continuation payoff there can be chosen to be no larger than what he can secure at that point by deviating and experimenting again, etc. The right-hand side is then obtained as the payoff from N consecutive unilateral deviations to experimentation (in fact, we have picked an upper bound, as the continuation payoff after this string of deviations need not be the maximum w_N). The left-hand side is the payoff from playing safe one period before setting $\alpha = 1$ and getting the maximum payoff w_N, a continuation strategy that is sequentially rational given that $\alpha = 1$ is enforceable at l by our hypothesis that $\alpha = 0$ is enforceable at l_{N-1}.

Plugging in the definition of ω_N, this inequality simplifies to

$$(\delta^2 - \delta^N)\omega_N \geq (\delta^2 - \delta^N)(l_N - m) + (1-\delta)(l-m),$$

which is always satisfied for beliefs $l \leq m$, *i.e.* below the myopic cutoff l^m (which coincides with the normalized payoff m).

To summarize, if $\alpha = 0$ can be enforced at the $N-1$ consecutive beliefs $l_1, \ldots, l_{N-1}$, with $l_N \geq l^*$ and $l \leq l^m$, then both $\alpha = 0$ and $\alpha = 1$ can be enforced at l. By induction, this implies that if we can find an interval of beliefs $[l_N, l)$ with $l_N \geq l^*$ for which $\alpha = 0$ can be enforced, then $\alpha = 0, 1$ can be enforced at all beliefs $l' \in (l, l^m)$.

Our second step is to establish that such an interval of beliefs exists. This second step involves itself three steps. First, we derive some "simple" equilibrium, which is a symmetric Markov equilibrium. Second, we will show that we can enforce $\alpha = 1$ on sufficiently (finitely) many consecutive values of beliefs building on this equilibrium; third, we show that this can be used to enforce $\alpha = 0$ as well.

It will be useful to distinguish beliefs according to whether they belong to the interval $[l^*, (1+\lambda_1\Delta)l^*), [(1+\lambda_1\Delta)l^*, (1+2\lambda_1\Delta)l^*), \ldots$ For $\tau \geq 0$, let $I_{\tau+1} = [(1+\tau\lambda_1\Delta)l^*, (1+(\tau+1)\lambda_1\Delta)l^*)$. For fixed Δ, every $l \geq l^*$ can be uniquely mapped into a pair $(x, \tau) \in [0,1) \times I\!N$ such that $l = (1 + \lambda_1(x+\tau)\Delta)l^*$, and we alternatively denote beliefs by such a pair. Note also that, for small enough $\Delta > 0$, one unsuccessful experiment takes a belief that belongs to the interval $I_{\tau+1}$ to (within $O(\Delta^2)$ of) the interval I_τ. (Recall that $\Lambda = \lambda_1\Delta + O(\Delta^2)$.)

Let us start with deriving a symmetric Markov equilibrium. Hence, because it is Markovian, $\omega_0 = \omega$ in our notation, that is, the continuation payoff when nobody experiments is equal to the payoff itself.

Rewriting the equations, using the risky arm gives the payoff[9]

$$\omega = l - (1-\delta)m - \delta(1-\Lambda)(1-\alpha\Lambda)^{N-1}l + \delta\sum_{K=0}^{N-1}\binom{N-1}{K}\alpha^K(1-\alpha)^{N-1-K}\omega_{K+1},$$

[9]To pull out the terms involving the belief l from the sum appearing in the definition of ω, use the fact that $\sum_{K=0}^{N-1}\binom{N-1}{K}\alpha^K(1-\alpha)^{N-1-K}(1-\Lambda)^K = (1-\alpha\Lambda)^N/(1-\alpha\Lambda)$.

while using the safe arm yields

$$\omega = \delta(1-(1-\alpha\Lambda)^{N-1})l + \delta(1-\alpha)^{N-1}\omega + \delta\sum_{K=1}^{N-1}\binom{N-1}{K}\alpha^K(1-\alpha)^{N-1-K}\omega_K.$$

In the Markov equilibrium we derive, players are indifferent between both actions, and so their payoffs are the same. Given any belief l or corresponding pair (τ, x), we conjecture an equilibrium in which $\alpha = a(\tau,x)\Delta^2 + O(\Delta^3)$, $\omega = b(\tau,x)\Delta^2 + O(\Delta^3)$, for some functions a, b of the pair (τ, x) only. Using the fact that $\Lambda = \lambda_1\Delta + O(\Delta^2), 1-\delta = r\Delta + O(\Delta^2)$, we replace this in the two payoff expressions, and take Taylor expressions to get, respectively,

$$0 = \left(rb(\tau,x) + \frac{\lambda_1 m}{\lambda_1 + r}(N-1)a(\tau,x)\right)\Delta^3 + O(\Delta^4).$$

and

$$0 = \left[b(\tau,x) - rm\lambda_1(\tau+x)\right]\Delta^2 + O(\Delta^3).$$

We then solve for $a(\tau,x)$, $b(\tau,x)$, to get

$$\alpha_- = \frac{r(\lambda_1+r)(x+\tau)}{N-1}\Delta^2 + O(\Delta^3),$$

with corresponding value

$$\omega_- = \lambda_1 mr(x+\tau)\Delta^2 + O(\Delta^3).$$

This being an induction on K, it must be verified that the expansion indeed holds at the lowest interval to which it is meant to hold, I_1, and this verification is immediate. (Note that this solution is actually continuous at the interval endpoints as well).[10]

We now turn to the second step and argue that we can find $N-1$ consecutive beliefs at which $\alpha = 1$ can be enforced. We will verify that incentives can be provided to do so, assuming that ω_- are the continuation values used by the players whether a player deviates or not from $\alpha = 1$. Assume that $N-1$ players choose $\alpha = 1$. Consider the remaining one. His incentive constraint to choose $\alpha = 1$ is

$$-(1-\delta)m + \delta\omega_N - \delta(1-\Lambda)^N l \geq -(1-\delta)l - \delta(1-\Lambda)^{N-1}l + \delta\omega_{N-1}, \tag{C.8}$$

where ω_N, ω_{N-1} are given by ω_- at l_N, l_{N-1}. The interpretation of both sides is as before, the payoff from abiding with the candidate equilibrium action vs. the payoff from deviating. Fixing l and the corresponding pair (τ, x), and assuming that $\tau \geq N-1$,[11] we insert our

[10] This is not the only solution to these equations; as mentioned in the text, there are intervals of beliefs for which multiple symmetric Markov equilibria exist in discrete time. It is easy to construct such equilibria in which $\alpha = 1$ and the initial belief is in (a subinterval of) $[l^*, (1+\lambda_1\Delta)l^*)$.

[11] Considering $\tau < N-1$ would lead to $\omega_N = 0$, so that the explicit formula for ω_- would not apply at l_N. Computations are then easier, and the result would hold as well.

formula for ω_-, as well as $\Lambda = \lambda_1 \Delta, 1 - \delta = r\Delta$. This gives

$$\tau \geq (N-1)\left(2 + \frac{\lambda_1}{\lambda_1 + r}\right) - x.$$

Hence, given any integer $N' \in I\!N$, $N' > 3(N-1)$, there exists $\bar{\Delta} > 0$ such that for every $\Delta \in (0, \bar{\Delta})$, $\alpha = 1$ is an equilibrium action at all beliefs $l = l^*(1+\tau\Delta)$, for $\tau = 3(N-1), \ldots, N'$ (we pick the factor 3 because $\lambda_1/(\lambda_1 + r) < 1$).

Fix $N-1$ consecutive values of beliefs such that they all belong to intervals I_τ with $\tau \geq 3(N-1)$ (say, $\tau \leq 4N$), and fix Δ for which the previous result holds, *i.e.* $\alpha = 1$ can be enforced at all these beliefs. We now turn to the third step, showing how $\alpha = 0$ can be enforced as well for these beliefs.

Suppose that players choose $\alpha = 0$. As a continuation payoff, we can use the payoff from playing $\alpha = 1$ in the following round, as we have seen that this action can be enforced at such a belief. This gives

$$\delta l + \delta(-(1-\delta)m - \delta(1-\Lambda)^N l + \delta\omega_-(l_N)).$$

(Note that the discounted continuation payoff is the left-hand side of (C.8).) By deviating from $\alpha = 0$, a player gets at most

$$l + (-(1-\delta)m - \delta(1-\Lambda)l + \delta\omega_-(l_1)).$$

Again inserting our formula for ω_-, this reduces to

$$\frac{mr(N-1)\lambda_1}{\lambda_1 + r}\Delta \geq 0.$$

Hence we can also enforce $\alpha = 0$ at all these beliefs. We can thus apply our induction argument: there exists $\bar{\Delta} > 0$ such that, for all $\Delta \in (0, \bar{\Delta})$, both $\alpha = 0, 1$ can be enforced at all beliefs $l \in (l^*(1 + 4N\Delta), l^m)$.

Note that we have not established that, for such a belief l, $\alpha = 1$ is enforced with a continuation in which $\alpha = 1$ is being played in the next round (at belief $l_N > l^*(1 + 4N\Delta)$). However, if $\alpha = 1$ can be enforced at belief l, it can be enforced when the continuation payoff at l_N is highest possible; in turn, this means that, as $\alpha = 1$ can be enforced at l_N, this continuation payoff is at least as large as the payoff from playing $\alpha = 1$ at l_N as well. By induction, this implies that the highest equilibrium payoff at l is at least as large as the one obtained by playing $\alpha = 1$ at all intermediate beliefs in $(l^*(1+4N\Delta), l)$ (followed by, say, the worst equilibrium payoff once beliefs below this range are reached).

Similarly, we have not argued that, at belief l, $\alpha = 0$ is enforced by a continuation equilibrium in which, if a player deviates and experiments unilaterally, his continuation payoff at l_1 is what he gets if he keeps on experimenting alone. However, because $\alpha = 0$ can be enforced at l_1, the lowest equilibrium payoff that can be used after a unilateral deviation at

l must be at least as low as what the player can get at l_1 from deviating unilaterally to risky again. By induction, this implies that the lowest equilibrium payoff at belief l is at least as low as the one obtained if a player experiments alone for all beliefs in the range $(l^*(1+4N\Delta), l)$ (followed by, say, the highest equilibrium payoff once beliefs below this interval are reached).

Note that, as $\Delta \to 0$, these bounds converge (uniformly in Δ) to the cooperative solution (restricted to no experimentation at and below $l = l^*$) and the single-agent payoff, respectively, which was to be shown. (This is immediate given that these values correspond to precisely the cooperative payoff (with N or 1 player) for a cut-off that is within a distance of order Δ of the cut-off l^*, with a continuation payoff at that cut-off which is itself within Δ times a constant of the safe payoff.)

This also immediately implies (as for the case $\lambda_0 > 0$) that for fixed $l > l^m$, both $\alpha = 0, 1$ can be enforced at all beliefs in $[l^m, l]$ for all $\Delta < \bar{\Delta}$, for some $\bar{\Delta} > 0$: the gain from a deviation is of order Δ, yet the difference in continuation payoffs (selecting as a continuation payoff a value close to the maximum if no player unilaterally defects, and close to the minimum if one does) is bounded away from 0, even as $\Delta \to 0$.[12] Hence, all conclusions extend: fix $l \in (l^*, \infty)$; for every $\epsilon > 0$, there exists $\bar{\Delta} > 0$ such that for all $\Delta < \bar{\Delta}$, the best SSE payoff starting at belief l is at least as much as the payoff from all players choosing $\alpha = 1$ at all beliefs in $(l^* + \epsilon, l)$ (using s as a lower bound on the continuation once the belief $l^* + \epsilon$ is reached); and the worst SSE payoff starting at belief l is no more than the payoff from a player whose opponents choose $\alpha = 1$ if and only if $l \in (l^*, l^* + \epsilon)$, and 0 otherwise.

The first part of the Proposition follows immediately, picking arbitrarily $p^\S \in (p_1^*, p^m)$ and $p^\sharp \in (p^m, 1)$. The second part follows from the fact that (i) $p_1^* < p_1^\Delta$, as noted, and (ii) for any $p \in [p_1^\Delta, p^\S]$, player i's payoff in any equilibrium is weakly lower than his best-reply payoff against $\kappa(p) = 1$ for all $p \in [p_1^*, p^\S]$, as easily follows from (C.7), the optimality equation for w.[13] ■

References

Abreu, D. (1986): "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, **39**, 195–225.

Abreu, D., D. Pearce and E. Stacchetti (1986): "Optimal Cartel Equilibria

[12]This obtains by contradiction. Suppose that for some $\Delta \in (0, \bar{\Delta})$, there is $\hat{l} \in [l^m, l]$ for which either $\alpha = 0$ or 1 cannot be enforced. Consider the infimum over such beliefs. Continuation payoffs can then be picked as desired, which is a contradiction as it shows that at this presumed infimum belief $\alpha = 0, 1$ can in fact be enforced.

[13]Consider the possibly random sequence of beliefs visited in an equilibrium. At each belief, a flow loss of either $-(1-\delta)m$ or $-(1-\delta)l$ is incurred. Note that the first loss is independent of the number of other players' experimenting, while the second is necessarily lower when at each round all other players experiment.

with Imperfect Monitoring," *Journal of Economic Theory*, **39**, 251–269.

Abreu, D., D. Pearce and E. Stacchetti (1993): "Renegotiation and Symmetry in Repeated Games," *Journal of Economic Theory*, **60**, 217–240.

Biais, B., T. Mariotti, G. Plantin and J.-C. Rochet (2007): "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, **74**, 345–390.

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

Cronshaw, M. B. and D.G. Luenberger (1994): "Strongly Symmetric Subgame Perfect Equilibria in Infinitely Repeated Games with Perfect Monitoring and Discounting," *Games and Economic Behavior*, **6**, 220–237.

Heidhues, P., S. Rady and P. Strack (2012): "Strategic Experimentation with Private Payoffs," SFB TR 15 Discussion Paper No. 387.

Keller, G. and S. Rady (2010): "Strategic Experimentation with Exponential Bandits," *Theoretical Economics*, **5**, 275–311.

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

Sadzik, T. and E. Stacchetti (2012): "Agency Models with Frequent Actions: A Quadratic Approximation Method," working paper, New York University.