Breakdowns*

Godfrey Keller† Sven Rady‡

This version: December 18, 2012

Abstract

We study a continuous-time game of strategic experimentation in which the players try to assess the failure rate of some new equipment or technology. Breakdowns occur at the jump times of a Poisson process whose unknown intensity is either high or low. In marked contrast to existing models, we find that the cooperative value function does not exhibit smooth pasting at the efficient cut-off belief. This finding extends to the boundaries between continuation and stopping regions in Markov perfect equilibria. We characterize the unique symmetric equilibrium, construct a class of asymmetric equilibria, and elucidate the impact of bad versus good Poisson news on equilibrium outcomes.

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

JEL Classification Numbers: C73, D83, O32.

*We would like to thank Christine Grün, Mike Ludkovski, Moritz Meyer-ter-Vehn, Philipp Strack and Bruno Strulovici for helpful discussions and suggestions, and the Studienzentrum Gerzensee for its hospitality during ESSET 2011 and 2012. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

†Department of Economics, University of Oxford, Manor Road Building, Oxford OX1 3UQ, UK; godfrey.keller@economics.ox.ac.uk.

‡Department of Economics and Hausdorff Center for Mathematics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany; rady@hcm.uni-bonn.de.

1 Introduction

The adoption of new technologies crucially hinges upon an assessment of the risks they might entail. It is very important, therefore, to obtain accurate information about the frequency of critical events, their severity, the size of the associated costs, etc. On a smaller scale, the introduction of a new production process or machine in a manufacturing plant will be based at least partly on the expected frequency with which the equipment fails and on the expected costs required to render it operational again. Similarly, a consumer's decision to acquire an innovative household appliance or some novel hardware will depend on an initial estimate of the product's reliability.

Once the new technology, process or equipment is in use, the assessment of its failure rate will be continually revised on the basis of one's own experience and, possibly, that of other users whose choices and results one may be able to observe. Each failure makes the users more pessimistic, and a string of such events may eventually lead them to abandon the exploration and switch to some alternative whose risks are better known.

We model this joint exploration process as a continuous-time game of strategic experimentation with identical two-armed bandits. The risky arm represents the 'machine' whose reliability is being explored. It imposes lump-sum costs at the jump times of a Poisson process; the arrival rate of these 'breakdowns' can be high or low, and is initially unknown.[1] The safe arm represents a machine with known costs per unit of time. If the risky machine is good, that is, if it fails at the lower rate, it is cheaper to maintain than the safe one; the opposite holds if the risky machine is bad.

We assume that the risky machines are either all good or all bad; conditional on this common quality, lump-sum costs arrive according to independent Poisson processes. The players can observe each other's choices and outcomes, so there is an informational externality among them. To gauge its effects, we characterize the efficient strategy profile and construct Markov perfect equilibria where the players' common assessment of the unknown failure rate serves as the state variable.

We first consider the case where a good risky machine never fails, so that any breakdown provides conclusive evidence of its being bad. In this case, efficient behaviour leads to a value function that is continuous and piecewise linear with a single kink at the boundary between the continuation and stopping regions. Thus, the cooperative value function does not exhibit smooth pasting at the efficient cut-off belief. This finding extends to the boundaries between continuation and stopping regions in Markov perfect equilibria of the experimentation game.

[1] While the failure rate is exogenous in our model, Biais, Mariotti, Rochet and Villeneuve (2010) consider a principal-agent problem in which this rate depends on the unobservable effort exerted by the agent.

With conclusive breakdowns, any such equilibrium leads to the efficient outcome on the *same* interval of initial beliefs. At beliefs optimistic enough to let the risky arm dominate the safe one in terms of expected current costs, in fact, the equilibrium path of play is clearly the efficient one, with all players sticking to the risky arm as long as there is no breakdown, and switching to the safe arm irrevocably as soon as a breakdown occurs. At somewhat less optimistic beliefs, this action profile remains compatible with equilibrium (and induces a common value function equal to the planner's solution) up to the point where a player whose opponents all play risky is just indifferent; the corresponding threshold belief depends on the number of players, but not on the precise structure of the equilibrium being played. Backward induction from this threshold allows us to construct equilibrium actions at more pessimistic beliefs and, despite the lack of smooth pasting, determine the boundary of the stopping region. We compute the unique symmetric Markov equilibrium, construct a class of asymmetric equilibria for two players, and indicate how this construction generalizes to an arbitrary number of players.

In the case where a good risky machine also fails occasionally, breakdowns provide inconclusive evidence of the true state of the world, and the belief held immediately after a breakdown may still be optimistic enough to continue using the risky machine. Put differently, whether it is optimal to use the risky machine at a given belief now depends on what would be the continuation payoff after a breakdown – this renders the analysis significantly more difficult. Efficient behaviour is still given by a cut-off strategy, but the optimal cut-off can no longer be computed in closed form. We show that it is uniquely determined by the requirement that the associated total expected cost function be continuous, that is, by value matching alone. We again establish existence of a unique symmetric Markov perfect equilibrium and characterize its properties, among them continuity and monotonicity of the equilibrium strategy. Finally, we briefly address the construction of asymmetric equilibria.

Our model of breakdowns is isomorphic to the setup considered by Keller and Rady (2010) except for the replacement of lump-sum *payoffs* (whose expected total net present value players want to maximize) with lump-sum *costs* (and the corresponding minimization objective); the special case of conclusive breakdowns corresponds to the setup with fully revealing 'breakthroughs' of Keller, Rady and Cripps (2005). One might have conjectured that a model of breakdowns (where news is bad) would lead to results that were just mirror images of those arising in an otherwise identical model of breakthroughs (where news is good), but this is not so.

Above all, the principle of smooth pasting does not apply here – value functions have a kink at the boundary between the continuation and stopping regions. The reason for this striking difference lies in the behaviour of the process of posterior beliefs when started at the boundary of the stopping region. Owing to the finite arrival rate of Poisson jumps, there will almost surely be no news event (breakthrough or breakdown, respectively) over the next instant. In the 'breakthroughs' case, no news is bad news, and the belief thus immediately enters the interior of the stopping region (in the terminology of the mathematical literature on optimal stopping, this means that the boundary is 'regular'); in the 'breakdowns' case, by contrast, no news is good news, and the belief moves away from the stopping region. The lack of smooth pasting at the efficient cut-off confirms the rule of thumb whereby the value function of a stopping problem is differentiable at a regular boundary, but not necessarily at an irregular one.[2]

In our framework, it is actually quite easy to understand why the cooperative value function cannot be differentiable at the efficient cut-off. At each point in time, the planner compares the expected informational benefit of using the risky arm with the expected (shared) cost increment relative to the safe arm. Both depend on the planner's belief, that is, the probability he assigns to the good state of the world. The informational benefit has two components, one capturing a gradual improvement in the overall outlook if no breakdown occurs, the other a discrete deterioration if a breakdown does occur; the former depends on the first derivative of the value function with respect to the belief, the latter on the difference between the continuation value at the belief held immediately after a breakdown and the value at the current belief. As is standard in this type of problem, the interior of the continuation region (where beliefs are more optimistic than the efficient cut-off) is characterized by the informational benefit exceeding the expected cost increment of using the risky arm, and the interior of the stopping region by the converse inequality; at the cut-off itself, benefits and costs are equal.

Now, the crucial insight is that in the interior of the stopping region, the informational benefit of experimentation is zero: in the absence of a breakdown, a planner 'deviating' to the risky arm would become slightly more optimistic, but then still not find it optimal to experiment; and if a breakdown did occur, the planner would find himself even 'deeper' in the stopping region than before and hence see no reason to experiment either. As a consequence, the benefit of experimentation must possess a jump discontinuity at the efficient cut-off, where the expected cost increment of using the risky arm is necessarily positive. The discrete-deterioration component of this benefit is continuous in the belief, however, so it must be the gradual-improvement component that jumps, implying a jump discontinuity in the first derivative of the value function at the cut-off.[3]

[2] See Peskir and Shiryaev (2006), especially Chapters IV.9 and VI.23.

[3] In the scenario with breakthroughs, the benefit of experimentation consists of a gradual-deterioration and a discrete-improvement component, and as we approach the efficient cut-off from the interior of the stopping region, the latter component is positive and increases monotonically, while the former component is zero. This makes it possible for the discrete-improvement component alone to balance the opportunity cost of experimentation at the efficient cut-off, so that the gradual-deterioration component – and hence the first derivative of the value function – is continuous there.

This argument carries over to any player who chooses a best response against opponents whose actions change continuously with the players' common posterior belief, the only difference being that unlike the social planner, each individual player compares the benefit of experimentation with the full expected cost increment of using the risky arm, not the shared one. Thus, the players' common value function in the symmetric equilibrium must have a kink at the threshold belief at which experimentation starts. And in an asymmetric equilibrium where experimentation starts with only one player playing risky, the value function of this player must have a kink at the corresponding threshold belief.[4]

Another difference between good and bad news is that in the scenario with breakdowns, the presence of other players always encourages experimentation in the sense that the equilibrium continuation region is larger than that of a single agent experimenting in isolation. While Keller and Rady (2010) established this encouragement effect for any Markov equilibrium of the experimentation game with inconclusive breakthroughs, Keller, Rady and Cripps (2005) had shown earlier that there is no such effect when breakthroughs provide conclusive evidence of the risky arm being good. With inconclusive good news, in fact, a player who experiments beyond the belief at which his opponents stop stands to bring them 'back into the game' if he has a breakthrough, and then benefit from their subsequent experiments. With conclusive breakthroughs, by contrast, those subsequent experiments are worthless because a successful 'pioneer' already knows everything there is to know about the quality of the risky arm; any such pioneer thus faces the same trade-off as a single agent experimenting in isolation, and no Markov perfect equilibrium can involve experimentation beyond the single-agent cut-off. It is noteworthy, therefore, that here we find an encouragement effect even in the case where a single arrival of bad news is conclusive. At second sight, this is fully in line with the finding in Keller and Rady (2010), however: the *absence* of bad news (whether conclusive or not) represents inconclusive good news, and this is what motivates players to venture beyond the single-agent cut-off belief.

Again for conclusive breakdowns, there are a number of further results that stand in marked contrast to Keller, Rady and Cripps (2005). To start with, the value of information to players intent on playing the symmetric MPE is no longer positive in the entire experimentation region: at relatively pessimistic prior beliefs in this region, the players would reject a free signal that induces a small lottery over posterior beliefs.[5] Moreover, it is no longer the case that the common outcome in the symmetric Markov perfect equilibrium is uniformly dominated by the average outcome in an asymmetric Markov perfect equilibrium that has each player use one arm exclusively at any given belief. In asymmetric equilibria, finally, players who free-ride on the information generated by others when

[4]Typically, it will have further kinks at more optimistic beliefs where an opponent's action changes discontinuously. This type of non-differentiability is familiar from Keller, Rady and Cripps (2005) and Keller and Rady (2010), and does not indicate a failure of the smooth-pasting principle.

[5]This finding carries over to inconclusive, but highly informative breakdowns.

the opportunity cost of experimentation is high do not benefit, but do worse than a player who experiments there. We shall discuss each of these findings in detail below.

In summary, the paper makes three main contributions. First, it shows that when players learn from occasional bad-news events, the principle of smooth pasting applies neither to the efficient benchmark nor to the Markov perfect equilibria of the experimentation game; for conclusive breakdowns, this point is made in an elementary fashion and by means of closed-form solutions. Second, the paper establishes existence and uniqueness of a symmetric Markov equilibrium in a situation where (as in the case of inconclusive bad news) neither smooth-pasting nor backward-induction techniques apply. Third, the paper carefully elucidates the impact of good versus bad Poisson news in a multi-agent bandit model.

After a discussion of the related literature, the paper proceeds as follows. Section 2 sets up the model. Section 3 studies the efficient benchmark and Markov perfect equilibria for conclusive breakdowns, Section 4 for inconclusive breakdowns. Section 5 concludes.

Related Literature

Our work is part of a growing literature on bandit-based games of learning and experimentation. Assuming that the cumulative payoff from the risky arm follows a Brownian motion with unknown drift, Bolton and Harris (1999) prove existence of a unique symmetric Markov perfect equilibrium and show that it exhibits the encouragement effect. Bolton and Harris (2000) characterize all Markov perfect equilibria of the undiscounted limit of their model. Keller, Rady and Cripps (2005) and Keller and Rady (2010) maintain the structure of the Bolton-Harris model, but replace Brownian payoffs with a compound Poisson process; the connection of the present paper with these two articles has already been discussed in detail.

Owing to their tractability, the learning dynamics associated with conclusive Poisson news have repeatedly been used as building blocks for richer models. Examples of the good-news variety are the models of R&D competition of Malueg and Tsutsui (1997) and Besanko and Wu (2012), and the (discrete-time) financial contracting model of Bergemann and Hege (1998, 2005). Décamps and Mariotti (2004) analyse a duopoly model of irreversible investment with a learning externality and a public background signal that produces conclusive bad news; since a firm stops learning once it is *optimistic* enough to invest, the stopping boundary is regular, however, and so the smooth-pasting principle applies.[6] Strulovici (2010) investigates how individual experimentation on a two-armed

[6]Bergemann and Hege (1998, 2005) allow for moral hazard in the allocation of funds, Décamps and Mariotti (2004) for privately observed investment costs. For further related work on strategic learning with private information, see Rosenberg, Solan and Vieille (2007), Moscarini and Squintani (2010), Acemoglu,

bandit interacts with collective decision making through voting. His benchmark model assumes conclusive good news; in an extension, he allows for bad news and alludes to the failure of smooth pasting. Klein and Rady (2011) allow the type of the risky arm to be negatively correlated across players, so that good news for one player is bad news for the other. They show that equilibrium value functions are discontinuous at the boundary between adjacent intervals that both are absorbing for the learning dynamics – such pairs of intervals do not exist in our model.

By their very nature, economic models of rational learning are closely related to models of sequential testing in mathematical statistics. This link is especially tight in our case. On the one hand, this is because the very first formulation of smooth-pasting conditions appears in Mikhalevich (1958), a paper dealing with the sequential testing of two simple hypotheses about the unknown drift parameter of an observed Brownian motion.[7] On the other hand, the analysis of the corresponding problem for a Poisson process in Peskir and Shiryaev (2000) provides the first example for the *failure* of smooth pasting. Observing a Poisson process whose unknown intensity can be either high or low, and revising beliefs in exactly the same way as in our model, the decision maker in that paper faces the combined task of stopping the process and deciding which of the two rates to 'accept'; his objective is to minimize a weighted sum of the observation costs and the probabilities of making an error of the first and second kind, respectively. The optimal continuation region lies in between two threshold levels for the probability assigned to the high intensity. In between Poisson jumps, this probability declines gradually, which makes the lower threshold a regular stopping boundary and the upper threshold an irregular one. Correspondingly, the optimal solution is found by imposing smooth pasting at the lower threshold, but only continuous pasting at the upper threshold. With a different objective function, the social optimum and best responses in the good-news scenario of Keller and Rady (2010) mirror the situation at the lower threshold, while their bad-news counterparts in the present paper mirror the situation at the upper threshold.

In the mathematical finance literature, several papers report a failure of smooth pasting in stopping problems where the underlying asset price follows a Lévy process without a Gaussian component; see Boyarchenko and Levendorskiĭ (2002) or Alili and Kyprianou

Bimpikis and Ozdaglar (2011), Bonatti and Hörner (2011), Hopenhayn and Squintani (2011), Murto and Välimäki (2011), and Heidhues, Rady and Strack (2012).

[7]Starting with Samuelson (1965), smooth-pasting conditions have been used extensively in (financial) economics to solve problems of optimal stopping or control of one-dimensional diffusions; see Dumas (1991), Dixit (1993), Dixit and Pindyck (1994) for classic references, and Strulovici and Szydlowski (2012) for a recent contribution. All these papers specify a diffusion coefficient that is bounded away from zero, which ensures regularity of stopping boundaries (and hence smooth pasting) even when the underlying process can jump; see Bayraktar and Xing (2012). A case in point is the single-agent bandit model of Cohen and Solan (2012), where payoffs on the risky arm are given by a Lévy process and posterior beliefs follow a jump diffusion.

(2005) for the valuation of an American option, and Gapeev and Kühn (2005) as well as Baurdoux, Kyprianou and Pardo (2011) for the zero-sum stopping game played by the holder and the issuer of a convertible bond. Outside this particular literature, we are aware of only one paper that identifies a lack of smooth pasting in an economically motivated setting. Ludkovski and Sircar (2012) study optimal resource exploration in a model where new discoveries occur according to a jump process whose intensity is given by the exploration effort. Under a monopolist's optimal policy, costly exploration takes place between two threshold levels for current reserves. As these reserves diminish in between discoveries, the upper threshold is an irregular stopping boundary, and the value function is not smooth there. The extension of the model to a Cournot duopoly with a 'green' second producer who has access to an inexhaustible but relatively expensive source is analysed numerically only. Our paper is thus the first in the economics literature to identify in a fully analytic fashion – and even in closed form when breakdowns are conclusive – a failure of smooth pasting in a continuous-time, non-zero-sum stochastic game.

Through the underlying theme of bad versus good news, our work is loosely linked to a set of papers which investigate the impact of different signal structures on the equilibria of dynamic games with imperfect public monitoring. Abreu, Milgrom and Pearce (1991) study a discrete-time repeated partnership game in which the players observe the realization of a Poisson process. They find that Poisson events conveying bad news, that is, an increased likelihood of 'cheating', lead to more efficient equilibria than Poisson events conveying good news; in particular, only the bad-news case admits a non-trivial limit equilibrium as the period length goes to zero. Among other results, Fudenberg and Levine (2007) confirm this finding for a repeated commitment game with a long-run and a short-run player. Faingold and Sannikov (2011) study continuous-time reputation dynamics in a game where a large player faces a population of small players. In the special case where the public signals about the large player's behaviour are driven by a Poisson process, the equilibrium is unique if Poisson events are good news; when they are bad news, multiple equilibria are possible. Clearly, the modelling frameworks in these papers, and the economic forces behind their results, are very different from ours.

In recent work that is closer to ours, Board and Meyer-ter-Vehn (2012) examine reputation building by a seller when product quality (which can be either high or low) is persistent and depends stochastically on past investments. Consumers learn about quality through Poisson signals; the probability that they assign to the high quality measures the seller's reputation, and constitutes a natural state variable for Markov perfect equilibria. With conclusive bad news, incentives to invest increase in reputation and there is a continuum of equilibria with divergent dynamics that lead, for the same reason as in Klein and Rady (2011), to a discontinuous value function. With conclusive good news, incentives to invest decrease in reputation, and there is a unique cut-off below which the seller invests, leading to ergodic reputation dynamics and a continuous value function; this type of equilibrium is also shown to exist for a large class of inconclusive Poisson news. Rather than solving the relevant Bellman equations, the authors analyse the seller's optimization problem by means of a path integral that represents the value of high quality. As a consequence, pasting principles are not of the essence here.

2 A Model of Stochastic Breakdowns

The set-up of the model is that of Keller and Rady (2010) except for the fact that here events occurring on the risky arm are bad news. The Bellman equations stated below follow from exactly the same arguments as in Keller, Rady and Cripps (2005); see also Davis (1993).

There are $N \geq 1$ players, each of them endowed with one unit of a perfectly divisible resource per unit of time, and each facing a two-armed bandit problem in continuous time. The risky arm R generates lump-sum costs which are independent draws from a time-invariant distribution on $]0, \infty[$ with known mean h. If a player allocates the fraction $k_t \in [0, 1]$ of her resource to R over an interval of time $[t, t + dt[$, the probability of such a breakdown on R at some point in the interval is $k_t \lambda_\theta \, dt$, where $\theta = 1$ if R is bad, $\theta = 0$ if R is good, and $\lambda_1 > \lambda_0 \geq 0$ are constants known to all players. Conditional on θ, the arrival of lump-sum costs is independent across players. The fraction $1 - k_t$ allocated to the safe arm S causes an expected cost of $(1 - k_t)s \, dt$, where $s > 0$ is a constant known to all players. Therefore, the overall expected cost increment conditional on θ is $[(1 - k_t)s + k_t \lambda_\theta h] \, dt$. We assume that $\lambda_0 h < s < \lambda_1 h$, so each player prefers R to S if R is good, and prefers S to R if R is bad.

Players start with a common prior belief about the unknown state of the world θ. Observing each other's actions and outcomes, they hold common posterior beliefs throughout time. With p_t denoting the subjective probability at time t that players assign to the risky arm being bad, a player's expected cost increment conditional on all available information is $[(1 - k_t)s + k_t \lambda(p_t)h] \, dt$ with

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

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

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

where the expectation is over the stochastic processes $\{k_t\}$ and $\{p_t\}$, and $r > 0$ is the common discount rate.

As long as no lump-sum cost arrives, the belief evolves smoothly with infinitesimal increment $dp_t = -K_t \Delta\lambda\, p_t(1-p_t)\,dt$ where $K_t = \sum_{n=1}^N k_{n,t}$ is the overall intensity of experimentation, and $\Delta\lambda = \lambda_1 - \lambda_0$. If any of the players incurs a lump-sum cost at time t, the belief jumps up from p_{t-} (the limit of beliefs held before the arrival of the lump-sum cost) to $p_t = j(p_{t-})$ where

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

Players are restricted to Markov strategies $k_n \colon [0,1] \to [0,1]$ with the left limit belief p_{t-} as the state variable, so that the action player n takes at time t is $k_n(p_{t-})$. We require each strategy to be left-continuous and piecewise Lipschitz-continuous. This ensures that each strategy profile $(k_1, k_2, \ldots, k_N)$ induces, for any prior p, a well-defined law of motion for players' common beliefs and well-defined total expected costs $u_n(p|k_1, k_2, \ldots, k_N)$ for each individual player. These costs are continuous in p on any interval where the overall intensity of experimentation is positive; jump discontinuities can occur at priors p where the intensity 'lifts off' from its lower bound, being zero at p and positive on $]p-\epsilon, p[$. Furthermore, u_n is once continuously differentiable in p on any interval where k_n and $K_{\neg n} = \sum_{\ell \neq n} k_\ell$ (the intensity of experimentation carried out by player n's opponents) are both continuous, and at least one of them is positive; otherwise, u_n can have a kink.[8]

A strategy k_n is a cut-off strategy if there is a belief $\hat{p}$ such that $k_n(p) = 1$ for all $p \leq \hat{p}$, and $k_n(p) = 0$ otherwise. As an example, consider an infinitely impatient agent who merely weighs the short-run cost from playing the safe arm, s, against the expected short-run cost from playing the risky arm, $\lambda(p)h$. Such an agent would optimally use the myopic cut-off belief

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

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

When the players act cooperatively so as to minimize the average total cost per player, their common value function U_N^* is concave and continuous. Concavity reflects a non-negative value of information, and implies continuity in the open unit interval. Continuity at the boundaries follows from the fact that $U_N^*(p)$ is bounded above by the myopic cost $\lambda(p)h \wedge s$ and bounded below by the full-information cost $ps + (1-p)\lambda_0 h$, both of which converge to $\lambda_0 h = U_N^*(0)$ as $p \to 0$, and to $s = U_N^*(1)$ as $p \to 1$.

Moreover, U_N^* solves the Bellman equation

$$u(p) = s + \min_{K \in [0,N]} K\left\{c(p)/N - b(p,u)\right\}, \tag{1}$$

[8] We shall see below that such kinks always occur in equilibrium, whereas jumps are ruled out.

where K is the intensity of experimentation,

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

is the expected current cost increase from playing R rather than S, and

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

is the expected learning benefit of playing R. In fact, the term $-K\Delta\lambda\, p(1-p)u'(p)/r$ in the Bellman equation (1) captures the marginal improvement in the players' outlook while they experiment without a breakdown, and the term $K\lambda(p)\,[u(j(p)) - u(p)]/r$ the discrete deterioration at the time of a breakdown. As infinitesimal changes of the belief are always downward, we say that a continuous function u solves the Bellman equation if its left-hand derivative exists on $]0, 1]$ and (1) holds on $]0, 1[$ when this left-hand derivative is used to compute $b(p, u)$. The cooperative value function U_N^* is the unique solution satisfying the boundary conditions $u(0) = \lambda_0 h$ and $u(1) = s$.

If the shared extra cost of playing R exceeds the full expected benefit, the collectively optimal choice is $K = 0$ (all agents use S exclusively), and the cooperative value function satisfies $u(p) = s$. Otherwise, $K = N$ is optimal (all agents use R exclusively), and $u(p) = s + c(p) - Nb(p, u) = \lambda(p)h - Nb(p, u)$.

When $N \geq 2$ players act non-cooperatively, a strategy k_n^* for player n is a best response against his opponents' strategies $k_1, \ldots, k_{n-1}, k_{n+1}, \ldots, k_N$ if

$$u_n(p|k_1, \ldots, k_{n-1}, k_n^*, k_{n+1}, \ldots, k_N) \leq u_n(p|k_1, \ldots, k_{n-1}, k_n, k_{n+1}, \ldots, k_N)$$

for all priors p and all strategies k_n. The value function from playing a best response is continuous. Continuity at the boundaries of the unit interval follows from the same upper and lower bounds as in the cooperative case. Continuity in the interior follows from the observation that at a belief p where the overall intensity of experimentation lifts off as described above, any jump discontinuity in u_n would contradict the optimality of the strategy player n uses. In fact, $u_n(p-) > u_n(p) = s$ would imply costs above s immediately to the left of p, while $u_n(p-) < u_n(p) = s$ would imply that player n could do better by not playing safe at p.

Moreover, a strategy k_n^* is a best response for player n if and only if the resulting value function u_n solves the Bellman equation

$$u_n(p) = s - K_{\neg n}(p)\, b(p, u_n) + \min_{k_n \in [0,1]} k_n \left\{c(p) - b(p, u_n)\right\}, \tag{2}$$

and $k_n^*(p)$ achieves the minimum on the right-hand side at each belief p. The benefit of experimentation $b(p, u_n)$ is then non-negative at all beliefs. In fact, there are three cases. If $u_n(p) = s$, then this is a global maximum, so we must have a left-hand derivative $u_n'(p) \geq 0$; as $u_n(j(p)) \leq s$, moreover, we find $b(p, u_n) \geq 0$. If $u_n(p) < s$ and $k_n^*(p) = 0$ is an optimal action, the Bellman equation (2) implies $u_n(p) = s - K_{\neg n}(p)\, b(p, u_n)$ and hence $K_{\neg n}(p)\, b(p, u_n) > 0$. If $u_n(p) < s$ and $k_n^*(p) = 1$ is optimal, the Bellman equation yields $u_n(p) = \lambda(p)h - [K_{\neg n}(p) + 1]\, b(p, u_n)$; as $u_n(p) \leq \lambda(p)h \wedge s \leq \lambda(p)h$, this in turn implies $[K_{\neg n}(p) + 1]\, b(p, u_n) \geq 0$.

If $c(p) > b(p, u_n)$, then the optimal action is $k_n^*(p) = 0$, and equation (2) implies $u_n(p) = s - K_{\neg n}(p)\, b(p, u_n) \geq s - K_{\neg n}(p)\, c(p)$, with a strict inequality if $K_{\neg n}(p) > 0$. If $c(p) = b(p, u_n)$, then $k_n^*(p)$ is arbitrary in $[0, 1]$, and $u_n(p) = s - K_{\neg n}(p)\, c(p)$. Finally, if $c(p) < b(p, u_n)$, then $k_n^*(p) = 1$, and $u_n(p) = s - [K_{\neg n}(p) + 1]\, b(p, u_n) + c(p) < s - K_{\neg n}(p)\, c(p)$. When player n uses a best response, therefore, his optimal action at a given belief p depends on whether his value $u_n(p)$ is above, at, or below the level $s - K_{\neg n}(p)\, c(p)$.

A Markov perfect equilibrium (MPE) is a profile of strategies that are mutually best responses.

3 Conclusive Breakdowns

Suppose that $\lambda_0 = 0$ so that a good risky machine never breaks down. Then, $j(p) = 1$ and $u(j(p)) = s$ for all $p > 0$, which simplifies the analysis considerably.

3.1 Cooperative Solution

Fix an initial belief $p_0 = p$ and consider the following strategy: all players use the risky arm until a breakdown occurs, at which point all players irrevocably switch to the safe arm. With prior probability $1 - p$, this strategy generates total costs of zero as the risky machine never fails.

With prior probability p, the machine will fail for the first time at some random time τ, the lump-sum cost is incurred and the players suffer a flow cost of s for evermore; thus, the total discounted cost per player will be $e^{-r\tau}\left(\frac{rh}{N} + s\right)$. Taking expectations first with respect to the exponentially distributed variable τ and then with respect to the unknown state of the world, we compute the expected total costs per player as $\frac{N\lambda_1}{r+N\lambda_1}\left(\frac{rh}{N} + s\right) p$. These costs are smaller than s if and only if p is below the threshold stated in the following proposition.

Proposition 1 (Cooperative solution, $\lambda_0 = 0$) *If breakdowns are conclusive, the N-agent cooperative solution has all players use the safe arm above the cut-off belief*

$$p_N^* = \frac{(r + N\lambda_1)s}{(rh + Ns)\lambda_1} > p^m,$$

and the risky arm below. The cooperative value function is continuous, non-decreasing and piecewise linear with a single concave kink at p_N^*.

Proof: If the players adopt the stated strategy, then each player's payoff is

$$U_N^*(p) = \frac{(rh + Ns)\lambda_1}{r + N\lambda_1}\, p$$

when $p \leq p_N^*$, and $U_N^*(p) = s$ otherwise, implying the stated properties. For $p \leq p_N^*$, we have $s \geq U_N^*(p) = s + c(p) - Nb(p, U_N^*)$ and thus $b(p, U_N^*) \geq \frac{c(p)}{N}$. For $p > p_N^*$, we have $b(p, U_N^*) = 0 < \frac{c(p)}{N}$ as $p > p^m$. So U_N^* solves the Bellman equation (1), and hence is the value function for the cooperative problem. At all beliefs, the actions specified in the proposition achieve the minimum in the Bellman equation, so this common strategy is optimal. ■

The linearity of U_N^* to the left of the cut-off p_N^* reflects the fact that the players' actions are frozen until the random time when, in the bad state of the world, the first breakdown resolves all uncertainty. The concave kink at p_N^* reflects a positive value of information around the cut-off.

In view of the fact that either $K = 0$ or $K = N$ is optimal at any given belief, the cooperative problem can be viewed as a simple stopping problem. As already mentioned in the introduction, the failure of smooth pasting at the cut-off p_N^* is then fully in line with this cut-off being an irregular stopping boundary for the process of posterior beliefs. What is more, the arguments underlying the proof of Proposition 1 allow us to explain in a very elementary fashion why there *cannot* be smooth pasting at the socially optimal cut-off. To the right of it, in fact, the benefit of experimentation $b(p, u)$ must be zero because, with $u(p) = u(j(p)) = s$ and $u'(p) = 0$, both the 'slide benefit' $\Delta\lambda\, p(1-p)u'(p)/r$ and the 'jump disbenefit' $\lambda(p)\,[s - u(p)]/r$ vanish. The latter is continuous in p, so for the benefit of experimentation to cover the shared cost increment $c(p)/N > 0$ at the cut-off, the slide disbenefit must be positive there, which requires a positive left-hand derivative.[9]

[9]In the good-news scenario of Keller, Rady and Cripps (2005), by contrast, there are a slide *disbenefit* and a jump *benefit*, and as we approach the optimal cut-off from within the stopping region, the jump benefit increases to the point where it alone suffices to cover the shared opportunity costs of experimentation, so that the slide disbenefit – and hence the derivative of the value function – can indeed be zero at the optimal cut-off.

Instead of smooth pasting, it is the 'principle of continuous pasting' that applies here: amongst all possible common cut-offs, the socially optimal one is uniquely pinned down by the requirement that the cooperative value function be continuous. If all players used a cut-off $\hat{p} > p_N^*$, for instance, the average total cost per player would satisfy $u(\hat{p}) > s = u(\hat{p}+)$, and vice versa for $\hat{p} < p_N^*$.

Finally, we note that the value function for the cooperative problem can be recast as

$$U_N^*(p) = \lambda_1 hp - \frac{N\lambda_1}{r + N\lambda_1}(\lambda_1 h - s)p$$

when $p \leq p_N^*$. The first term, $\lambda_1 hp$, is the expected cost of committing to the risky arm, while the second term captures the option value of being able to change to the safe arm after the arrival of bad news.

3.2 Symmetric Equilibrium

The players' value functions in any MPE lie in the region of the (p, u)-plane below the graph of the myopic cost function, $\lambda_1 hp \wedge s$, and indeed below U_1^*. Define

$$\mathcal{D}_{K_{\neg n}} = \{(p, u) \in [0, 1] \times \mathbb{R}_+ : \; u = s - K_{\neg n}\, c(p)\}.$$

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

By the characterization of best responses in Section 2, the efficient actions described in Proposition 1 are mutually best responses whenever the graph of the cooperative value function U_N^* is weakly below the diagonal $\mathcal{D}_{N-1}$, that is, at beliefs no higher than

$$p_N^\dagger = \frac{(r + N\lambda_1)s}{[rh + s + (N-1)\lambda_1 h]\lambda_1}.$$

Recall that in this region U_N^* satisfies the ordinary differential equation (henceforth ODE) $u(p) = s - Nb(p, u) + c(p)$, so when U_N^* meets $\mathcal{D}_{N-1}$ we have $s - (N-1)c(p_N^\dagger) = s - Nb(p_N^\dagger, U_N^*) + c(p_N^\dagger)$, and its slope there is given by $b(p_N^\dagger, U_N^*) = c(p_N^\dagger)$.

The characterization of best responses in Section 2 further entails that in a symmetric MPE with common value function u, there are three cases: either all players use the safe arm exclusively and $u(p) = s$; or they all choose the interior allocation $k(p) = \frac{s - u(p)}{(N-1)\,c(p)}$ and u satisfies $b(p, u) = c(p)$ with $s - (N-1)\,c(p) < u(p) < s$; or they use the risky arm exclusively and $u(p) = s - Nb(p, u) + c(p) \leq s - (N-1)\,c(p)$. We know already that the latter case arises if and only if $p \leq p_N^\dagger$. Given this threshold, backward induction yields the following result.

Proposition 2 (Symmetric MPE, $\lambda_0 = 0$) *If breakdowns are conclusive, the N-player experimentation game has a unique symmetric Markov perfect equilibrium with the common posterior belief as the state variable. The equilibrium strategy is continuous and non-increasing, and has all players use the risky arm exclusively for $p \leq p_N^\dagger$. In addition, there is a threshold belief $\tilde{p}_N > p_N^\dagger$ with $p_1^* < \tilde{p}_N < p_N^*$ such that the players choose an interior allocation for $p_N^\dagger < p < \tilde{p}_N$ and use the safe arm exclusively for $p \geq \tilde{p}_N$. The equilibrium value function is continuous, strictly increasing on $[0, \tilde{p}_N]$, and once continuously differentiable except for a concave kink at $\tilde{p}_N$. On $[0, p_N^\dagger]$, it coincides with the cooperative value function U_N^*; on $[p_N^\dagger, \tilde{p}_N]$, it is strictly convex.*

Proof: As $u(j(p)) = s$, the indifference condition $c(p) = b(p, u)$ reduces to an ODE with the general solution

$$w(p) = rh + s - \frac{rs}{\lambda_1} + \frac{rs}{\lambda_1}(1-p)\ln\frac{1-p}{p} + C(1-p). \tag{3}$$

Choosing the constant C so that $w(p_N^\dagger) = s - (N-1)c(p_N^\dagger)$, we obtain a convex and increasing function W_N for $p \geq p_N^\dagger$ with $W_N(p_N^\dagger) = U_N^*(p_N^\dagger)$, and it follows from value matching together with $b(p_N^\dagger, W_N) = c(p_N^\dagger) = b(p_N^\dagger, U_N^*)$ that $W_N'(p_N^\dagger) = (U_N^*)'(p_N^\dagger)$. Let $\tilde{p}_N$ be the belief at which this function W_N reaches the cost level s, and define the Lipschitz-continuous strategy

$$k(p) = \begin{cases} 1 & \text{if } p \leq p_N^\dagger, \\ \frac{s - W_N(p)}{(N-1)\,c(p)} & \text{if } p_N^\dagger < p < \tilde{p}_N, \\ 0 & \text{if } p \geq \tilde{p}_N. \end{cases}$$

The function

$$u(p) = \begin{cases} U_N^*(p) & \text{if } p \leq p_N^\dagger, \\ W_N(p) & \text{if } p_N^\dagger < p < \tilde{p}_N, \\ s & \text{if } p \geq \tilde{p}_N \end{cases}$$

has the stated properties and satisfies the Bellman equation

$$u(p) = s - (N-1)k(p)\,b(p,u) + \min_{k \in [0,1]} k\,\{c(p) - b(p,u)\}$$

on $[0,1]$, with the minimum on the right-hand side achieved at $k^* = k(p)$. This proves that all players using the above strategy constitutes a symmetric MPE. Uniqueness follows from continuity of the equilibrium value function, the fact that it necessarily coincides with U_N^* on $[0, p_N^\dagger]$, and the fact that it cannot exceed the safe cost level s.

It remains to show that $\tilde{p}_N > p_1^*$. Since in any equilibrium, each player must be at least as well off as in the single-agent solution, we cannot have $\tilde{p}_N < p_1^*$. Suppose, therefore, that $\tilde{p}_N = p_1^*$. Then the equilibrium value function u and the single-agent value function

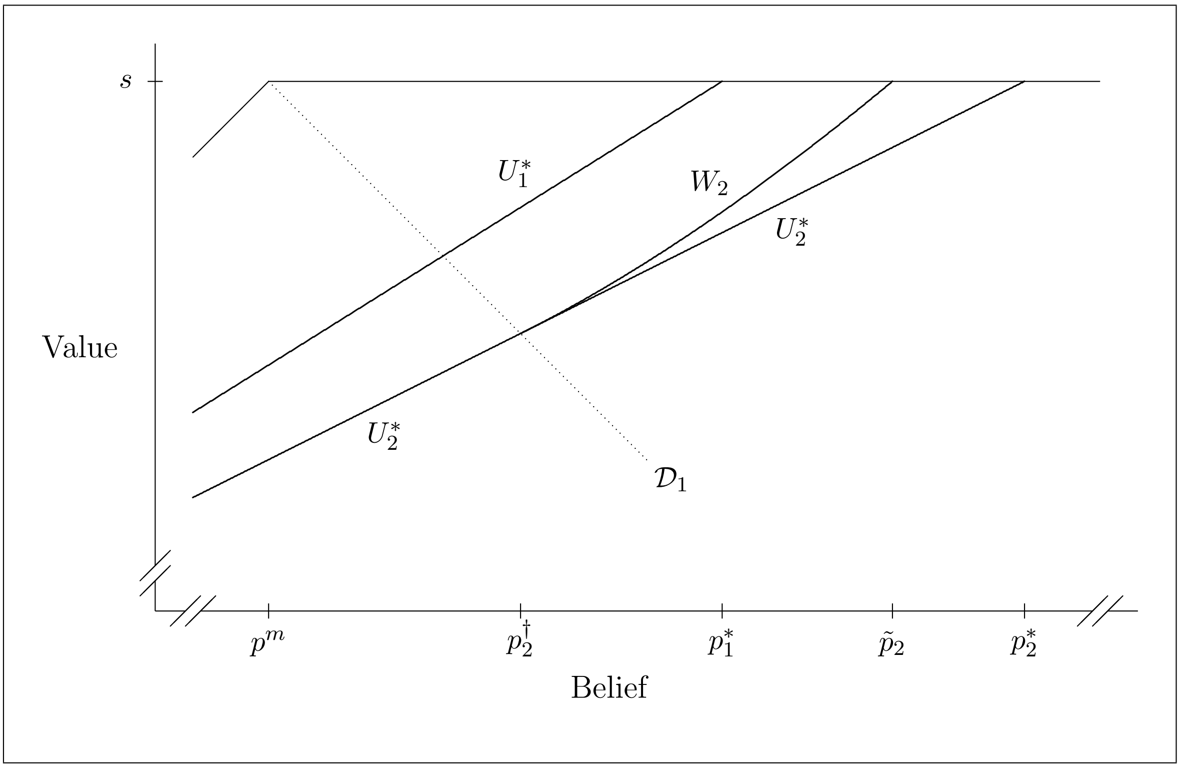

Figure 1: Value functions of a single agent, a two-person cooperative, and the two-player symmetric equilibrium ($\lambda_0 = 0$)

U_1^* satisfy $\Delta\lambda\, p_1^*(1-p_1^*)u'(p_1^*)/r = b(p_1^*, u) = c(p_1^*) = b(p_1^*, U_1^*) = \Delta\lambda\, p_1^*(1-p_1^*)(U_1^*)'(p_1^*)/r$ and hence $u'(p_1^*) = (U_1^*)'(p_1^*)$. Immediately to the left of p_1^*, strict convexity of u and linearity of U_1^* then imply $u > U_1^*$. This is impossible. ∎

For $N = 2$, the symmetric MPE is illustrated in Figure 1.

It is remarkable that even though each player optimizes against continuous behaviour of his opponents, and uses a continuous strategy himself, the resulting value function is not differentiable at the belief where experimentation with the risky arm 'takes off'. This lack of smooth pasting at the threshold belief $\tilde{p}_N$ can be explained in exactly the same way as in the cooperative problem. To the right of the threshold, both the slide benefit and the jump disbenefit of experimentation are zero. At the threshold, the slide benefit must be positive so as to cover the cost increment $c(\tilde{p}_N) > 0$, and this again requires a positive left-hand derivative.

That there is smooth pasting at the diagonal $\mathcal{D}_{N-1}$ can also be understood in terms of the benefits and costs of experimentation. This diagonal has been constructed as the locus of all pairs of a belief and a continuation value such that, when playing a best response against $N-1$ opponents who use the risky arm exclusively, the N^{th} player is indifferent between all possible intensities of experimentation. Thus, the benefit of experimentation exactly offsets its cost at $p_N^\dagger$ – just as it does to the right of this threshold, where all players use an interior allocation. Given that the jump disbenefit and the cost are continuous in beliefs, therefore, the slide disbenefit – and hence the first derivative of the equilibrium value function – must also be continuous at $p_N^\dagger$.

The equilibrium value function is strictly *convex* over the range of beliefs $]p_N^\dagger, \tilde{p}_N[$ associated with interior allocations (see the curve labelled W_2 in Figure 1). Starting from a prior p in this range, players who are intent on playing the symmetric MPE would thus reject any free signal about the unknown state of the world that induces a lottery (centred at p) over beliefs in $]p_N^\dagger, \tilde{p}_N[$. This negative value of information 'in the small' conforms to the familiar observation in multi-agent settings that the positive effect of additional information on one's own optimization can be overcome by the adverse effect of the concomitant change in the other agents' behaviour.[10] There is no contradiction, however, with the non-negative value of information 'in the large' that manifests itself in the globally non-negative benefit of experimentation $b(p, u_n)$ along any best response. In fact, the observation of a risky arm at an intensity $k > 0$ and over a length of time $\Delta > 0$ leads to a 'non-local' binary lottery with possible outcomes $p' = \frac{pe^{-k\Delta}}{1-p+pe^{-k\Delta}}$ and 1. And as can be seen in Figure 1, the straight line joining the point $(p', W_2(p'))$ with the point $(1, \lambda_1 h)$ is everywhere below the graph of the symmetric equilibrium value function, so that observing the risky arm indeed lowers total costs on average.

As $\tilde{p}_N > p_1^*$, the equilibrium exhibits an encouragement effect in the sense that it features experimentation on a strictly larger set of beliefs than would be optimal for a single agent experimenting in isolation. This effect is well-known from the unique symmetric MPE in the Brownian model of Bolton and Harris (1999) and in the Poisson model with inconclusive good news of Keller and Rady (2010). There, intuitively, each player is willing to experiment beyond the single-agent cut-off because any good news thus obtained makes all players more optimistic and increases the overall intensity of experimentation to everyone's benefit. When good news is conclusive, however, this reasoning breaks down: a player who experiences a breakthrough becomes certain of the good state of the world, and hence cannot learn anything from the opponents' subsequent use of the risky arm. In Keller, Rady and Cripps (2005), equilibrium experimentation thus stops at the single-agent cut-off p_1^*. The case of conclusive bad news is strikingly different in this respect, therefore. We can easily reconcile this finding with the above intuition, however, by noting that in the model with conclusive bad news, the *absence* of a breakdown represents inconclusive good news, and it is the prospect of generating this kind of good news that gives the players an incentive to experiment beyond p_1^*.

[10] The scenario with conclusive good news is different in this regard. In the symmetric MPE of Keller, Rady and Cripps (2005), the value of information 'in the small' is always non-negative, and positive in the entire experimentation region.

3.3 Asymmetric Equilibria and Welfare Properties

For conclusive good news, Keller, Rady and Cripps (2005) show that the symmetric MPE is dominated, in terms of average performance per player, by any asymmetric MPE in *simple* strategies; by definition, these are strategies which take values in $\{0, 1\}$ only and hence prescribe exclusive use of an arm at any given belief. Such equilibria have players take turns using the risky arm at beliefs slightly more optimistic than the single-agent cut-off where – owing to the lack of an encouragement effect – all experimentation stops. This keeps the intensity of experimentation bounded away from zero as the belief approaches the single-agent cut-off, whereas the symmetric MPE would see that intensity fall to zero so rapidly that the single-agent cut-off is actually not reached in finite time. A higher intensity of experimentation at relatively pessimistic beliefs implies better performance there, and this improvement 'ripples up' to more optimistic beliefs by backward induction. The inefficiency of the symmetric MPE close to the single-agent cut-off is actually so severe that, even though it might specify a higher aggregate intensity than a simple MPE over some range of more optimistic beliefs, its average performance remains worse there.[11]

We shall show that this unambiguous welfare comparison does not carry over to the scenario with conclusive bad news. The basis for this finding as well as for the construction of asymmetric equilibria is the observation that with conclusive breakdowns, *any* Markov perfect equilibrium of the N-player experimentation game coincides with the co-operative solution at all beliefs $p \leq p_N^\dagger$. In fact, once the value functions of all the players have crossed $\mathcal{D}_{N-1}$ from above, the profile of players' best responses is for them all to use the risky arm until a breakdown occurs. Therefore, below $\mathcal{D}_{N-1}$, each player's equilibrium cost function coincides with the cooperative value function U_N^* from Proposition 1. Asymmetric equilibria can thus also be constructed by backward induction from $p_N^\dagger$, and perform neither better nor worse than the symmetric MPE to the left of this threshold.

Proposition 3 (Welfare comparison, $\lambda_0 = 0$) *When breakdowns are conclusive, total costs per player immediately to the right of the threshold belief $p_N^\dagger$ are lower in the symmetric Markov perfect equilibrium of the N-player experimentation game than in any equilibrium in simple strategies.*

Proof: Consider an equilibrium in simple strategies with average cost function $\bar{u}$. Immediately to the right of $p_N^\dagger$, this function satisfies $\bar{u}(p) = s + K\{c(p)/N - b(p, \bar{u})\}$ for some $K \in \{1, 2, \ldots, N-1\}$. Recalling the function W_N defined in the proof of Proposition 2, we have $\bar{u}(p_N^\dagger) = s - (N-1)c(p_N^\dagger) = W_N(p_N^\dagger)$ and hence

$$b(p_N^\dagger, \bar{u}) = \frac{c(p_N^\dagger)}{N} + \frac{s - \bar{u}(p_N^\dagger)}{K} = \left[\frac{1}{N} + \frac{N-1}{K}\right] c(p_N^\dagger) > c(p_N^\dagger) = b(p_N^\dagger, W_N).$$

[11] There could be a range of beliefs, for example, where the simple MPE specifies the intensity $K = N-1$ while the symmetric MPE has $N - 1 < K < N$.

This implies $\bar{u}'(p_N^\dagger) > W_N'(p_N^\dagger)$. ■

In general, the symmetric MPE does not imply lower average costs all the way up to the belief $\tilde{p}_N$ at which experimentation takes off. With a view towards providing a counterexample, we note that when a player is volunteering to experiment with $K-1$ others, the Bellman equation (equation (2) with $K_{\neg n}(p) = K-1$ and $k_n = 1$) gives rise to the ODE $u(p) = s - Kb(p,u) + c(p)$ whose general solution is

$$v_K(p) = U_K^*(p) + C_v\,(1-p)\left(\frac{1-p}{p}\right)^{r/K\lambda_1}$$

with some constant of integration C_v. When a player is free-riding on the experimentation of K others, the Bellman equation (equation (2) with $K_{\neg n}(p) = K$ and $k_n = 0$) gives rise to the ODE $u(p) = s - Kb(p,u)$ whose general solution is

$$f_K(p) = s + C_f\,(1-p)\left(\frac{1-p}{p}\right)^{r/K\lambda_1}$$

with a constant C_f. Inspection of the ODEs for v_K and f_K shows that these functions are increasing whenever they are below the myopic cost function; their second derivative has the same sign as the respective constant of integration.

If $N = 2$, then, as noted above, both players play risky below and to the left of $\mathcal{D}_1$; above and to the right of $\mathcal{D}_1$, safe and risky are mutual best responses as long as the cost function of at least one player (and hence the average cost function $\bar{u}$) is below the level s. As $\bar{u}$ is increasing over the corresponding range of beliefs, there exists a threshold $\bar{p}_{2,1}$ with $p_2^\dagger < \bar{p}_{2,1} < p_2^*$ such that in any simple Markov perfect equilibrium of the two-player experimentation game, both players play risky when $p \le p_2^\dagger$, one of the two players is playing risky and the other safe when $p_2^\dagger < p \le \bar{p}_{2,1}$, and both are playing safe when $p > \bar{p}_{2,1}$. The assignment of roles within the interval $]p_2^\dagger, \bar{p}_{2,1}]$ is arbitrary. Figure 2 illustrates the assignment that leads to the most inequitable cost functions, with player A, the first volunteer, subsequently free-riding over the largest possible interval of beliefs.[12] More equitable value functions emerge simply by exchanging the roles of free-rider and volunteer more often, the only constraint being that once one of the value functions is at

[12] In the labelling of value functions to the right of $\mathcal{D}_1$ in Figure 2, the first subscript refers to the aggregate intensity of experimentation, $K = 1$, the second to the identity of the player. Note that when player B free-rides, he has a value function identical to s (which is trivially of the form f_1). Intuitively, player B will only ever switch from the safe to the risky arm if player A observes no breakdown while he acts as the first volunteer. As the burden of experimentation is subsequently borne by player B himself, player A's experimentation has indeed no option value for player B. This implies that player B is worse off than A, despite his free-riding when the costs of experimentation are high and experimenting when they are low.

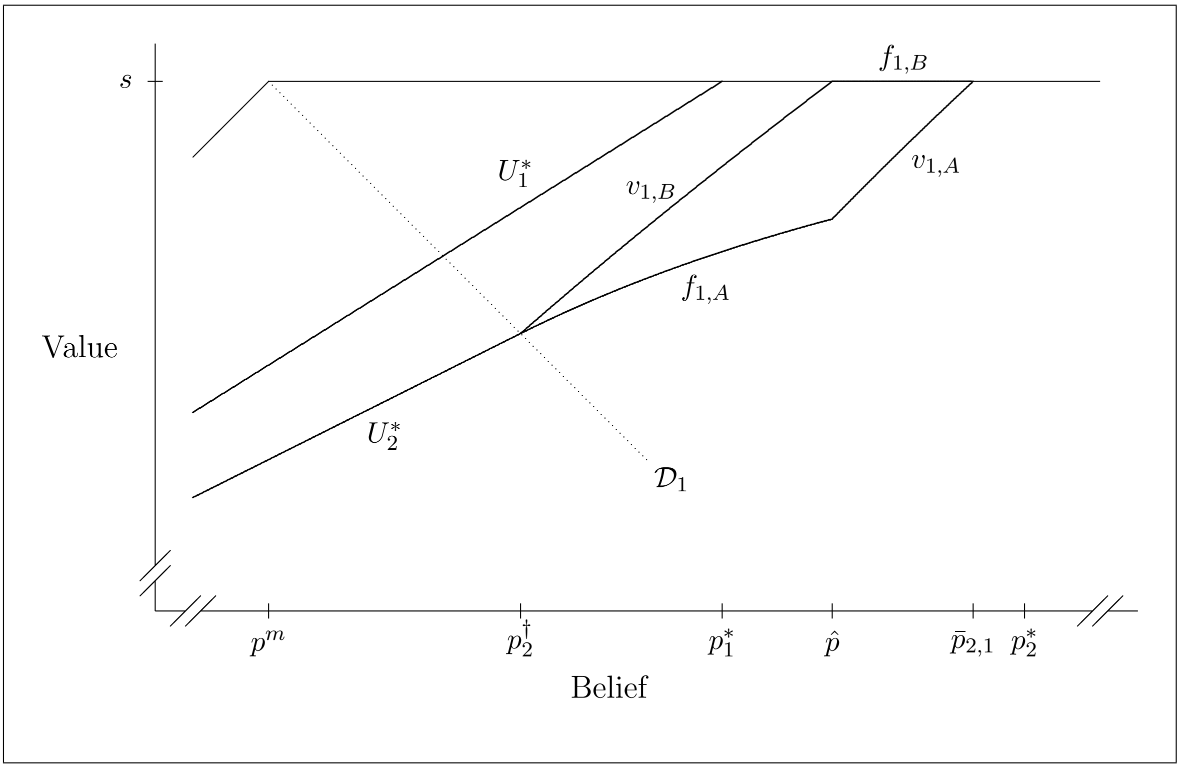

Figure 2: Value functions in a two-player asymmetric equilibrium ($\lambda_0 = 0$)

the level s, that player is assigned the safe role for all higher beliefs, and the other plays risky until their value function is also at the level s.

The numerical solutions (using parameter values $r = 1, s = 2, h = 8, \lambda_1 = 1$) illustrated in Figures 1 and 2 show that it is possible to have $\tilde{p}_2 < \bar{p}_{2,1}$; since $W_2(\tilde{p}_2) = s = \bar{u}(\bar{p}_{2,1})$, this implies that $W_2(p) > \bar{u}(p)$ immediately to the left of $\bar{p}_{2,1}$. In general, therefore, equilibria in simple strategies and the symmetric MPE cannot be compared in terms of aggregate welfare: while the latter performs better at relatively optimistic beliefs, the former have the potential to let experimentation take off earlier.

The above construction of a most inequitable equilibrium in simple strategies generalises to games involving more than two players, along the lines of Keller, Rady and Cripps (2005), Section 6. Figure 3 shows the assignment of actions in the most inequitable MPE for $N = 3$. (At and to the left of $\bar{p}_{N,K}$ at least K of the N players are playing risky; $\bar{p}_{N,N}$ can be identified with $p_N^\dagger$.) Unlike the two-player situation, the belief above which there is no experimentation is endogenously determined by how the burden of experimentation is shared in the interval to the right of where all agents play risky. In the situation depicted in Figure 3, for example, by changing roles between the free-rider and two volunteers for beliefs above $p_3^\dagger$ more frequently, we can increase $\bar{p}_{3,2}$ (the switch where aggregate exper-

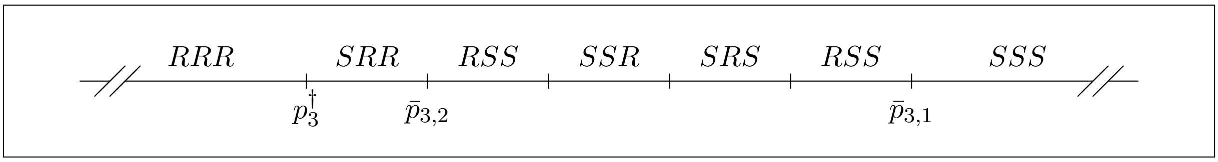

Figure 3: Action assignments in a three-player asymmetric equilibrium ($\lambda_0 = 0$)

imentation moves up from 1 to 2), with the concomitant increase in $\bar{p}_{3,1}$ (the threshold between no experimentation and some).

Exactly like the symmetric MPE, finally, any equilibrium in simple strategies exhibits a failure of smooth pasting at the boundary of the experimentation region as well as the encouragement effect, irrespectively of the number of players. In fact, if player A is the one to play risky on $]\hat{p}, \bar{p}_{N,1}]$, and player B the one to take over at $\hat{p}$, then the requirement that $b(\bar{p}_{N,1}, v_{1,A}) \geq c(\bar{p}_{N,1})$ implies $v'_{1,A}(\bar{p}_{N,1}) > 0$ by the same argument as before.[13] As player B has a value equal to s while free-riding, and this value cannot exceed U_1^*, moreover, we must have $p_1^* \leq \hat{p}$ and hence $p_1^* < \bar{p}_{N,1}$. Clearly, these statements remain true in any non-simple asymmetric equilibrium that has the players take turns playing risky immediately to the left of the threshold belief at which experimentation takes off.

4 Inconclusive Breakdowns

Now suppose that $\lambda_0 > 0$, so that even a good machine breaks down occasionally. In this case, $j(p) < 1$ for all $p < 1$, and the benefit of an experiment, $b(p, u)$, depends on the post-jump value $u(j(p))$.

We first show that the cooperative solution is again achieved by a common cut-off strategy, and that the cut-off is uniquely determined by continuous pasting. We then establish existence of a unique symmetric Markov perfect equilibrium (which again exhibits a failure of smooth pasting) and briefly address the problem of constructing asymmetric equilibria, which is considerably more difficult than in the case of conclusive breakdowns.

[13]Kinks at more optimistic beliefs are caused by discontinuities in the intensity of experimentation carried out by a player's opponents and, as such, have nothing to do with a failure of the smooth-pasting principle. In Figure 2, this remark concerns player A at belief $\hat{p}$, and player B at $p_2^\dagger$.

4.1 Cooperative Solution

Suppose that all players behave as in the cooperative solution for conclusive breakdowns, and switch to the safe arm as soon as one of them observes a breakdown. Then the expected total costs per player as a function of the initial belief are

$$\left(\frac{rh}{N} + s\right)\left[\frac{N\lambda_1}{r + N\lambda_1}\,p + \frac{N\lambda_0}{r + N\lambda_0}\,(1-p)\right]$$

by a straightforward generalisation of the computation leading up to Proposition 1. These costs are smaller than s if and only if p is below the threshold

$$\bar{p}_N = \frac{(r + N\lambda_1)(s - \lambda_0 h)}{(rh + Ns)\Delta\lambda} = \frac{(r + N\lambda_1)h}{(rh + Ns)}\,p^m > p^m.$$

At very optimistic beliefs, switching to the safe arm after a single (inconclusive) breakdown is clearly suboptimal, so the cooperative solution achieves a lower expected total cost than the one stated above. We should expect, therefore, that the cooperative will be willing to experiment to the right of $\bar{p}_N$. Our next proposition confirms this; its proof shows that the optimal cut-off is uniquely determined by continuous pasting.[14]

Proposition 4 (Cooperative solution, $\lambda_0 > 0$) *If breakdowns are inconclusive, the N-agent cooperative solution has all players use the safe arm above a unique cut-off belief $p_N^* > \bar{p}_N$, and the risky arm below. The cooperative value function is continuous, concave and non-decreasing; except for a kink at p_N^*, it is once continuously differentiable.*

Proof: Continuity and concavity of the cooperative value function have already been established in Section 2.

For arbitrary but fixed $\hat{p}$ in the open unit interval, consider the following profile of cut-off strategies: all players use the safe arm whenever $p > \hat{p}$, and the risky arm otherwise. Let $u_{\hat{p}}$ denote the players' corresponding common payoff function. For the common cut-off $\hat{p}$ to be collectively optimal, $u_{\hat{p}}$ must be continuous at $\hat{p}$. We wish to show that a unique such $\hat{p}$ exists and that the corresponding strategy profile solves the cooperative problem.

The mapping $\hat{p} \mapsto u_{\hat{p}}(\hat{p})$ is continuous with limit $\lambda_1 h$ as $\hat{p}$ tends to 1. For $\hat{p} = p^m$, we have

$$u_{p^m}(p^m) = \mathrm{E}\left[\int_0^\infty r\,e^{-r\,t}\,\min\left\{s, \lambda(p_t)h\right\}\,dt\right]$$

[14] In the statement of this proposition, we use the same notation for the optimal cut-off as in Proposition 1, although these cut-offs are not identical, of course. More precisely, p_N^* should be thought of as a function of λ_0 (holding all other model parameters fixed), with $\lambda_0 = 0$ leading to the expression given in Proposition 1. The same remark applies to the threshold beliefs $p_N^\dagger$ and $\tilde{p}_N$ in Proposition 2 and its counterpart, Proposition 5 below.

where the expectation is taken over the process $\{p_t\}$ induced by the given strategy profile for initial belief $p_0 = p^m$. As $\min\{s, \lambda(p)h\} < s + \Delta\lambda\, h\,(p - p^m)/2$ for $p \neq p^m$, the martingale property of posterior beliefs implies $u_{p^m}(p^m) < s$. By the intermediate-value theorem, therefore, there exists a $\hat{p}$ with $p^m < \hat{p} < 1$ such that $u_{\hat{p}}(\hat{p}) = s$; let p_N^* denote the smallest such belief, and u^* the corresponding common payoff function.

For $p \leq p_N^*$, we have $s \geq u^*(p) = s + c(p) - Nb(p, u^*)$ and thus $b(p, u^*) \geq \frac{c(p)}{N}$. For $p > p_N^*$, we have $b(p, u^*) = 0 < \frac{c(p)}{N}$ as $p > p^m$. So u^* solves the Bellman equation (1), with the minimum being achieved by the actions specified in the proposition. Thus, $u^* = U_N^*$, the cooperative value function, and the stated common strategy is optimal.

If there were another cut-off $\hat{p} > p_N^*$ such that $u_{\hat{p}}(\hat{p}) = s$, the corresponding payoff function would also coincide with the value function by the arguments just given. As this is impossible, p_N^* is uniquely pinned down by the continuous-pasting condition $u_{\hat{p}}(\hat{p}) = s$.

On $[0, p_N^*]$, the intensity of experimentation equals N, so u_N^* is continuously differentiable in this interval by what was said in Section 2. As $U_N^* = s$ on $[p_N^*, 1]$, moreover, concavity implies that U_N^* increases on $[0, p_N^*]$.

To establish that there is a kink at p_N^*, we note that on the interval $]j^{-1}(p_N^*), p_N^*[$, U_N^* solves the ODE

$$\Delta\lambda\, p(1-p)u'(p) + \left[\tfrac{r}{N} + \lambda(p)\right] u(p) = \lambda(p)\left[\tfrac{r}{N}h + s\right] \tag{4}$$

which is obtained from the identity $u(p) = s + N\{c(p)/N - b(p, u)\}$ by setting $u(j(p)) = s$ in $b(p, u)$ and rearranging. Letting p tend to p_N^* from below, we see that $u'(p_N^*-)$ has the same sign as $\lambda(p_N^*)\left[\frac{r}{N}h + s\right] - \left[\frac{r}{N} + \lambda(p_N^*)\right] s = \frac{r}{N}c(p_N^*)$. This is positive because $p_N^* > p^m$.

The general solution to (4) is

$$u(p) = \left(\frac{rh}{N} + s\right)\left[\frac{N\lambda_1}{r + N\lambda_1}\, p + \frac{N\lambda_0}{r + N\lambda_0}\,(1-p)\right] + C\,(1-p)\left(\frac{1-p}{p}\right)^{(r+N\lambda_0)/N\Delta\lambda} \tag{5}$$

where C is a constant of integration. As the value function is at least weakly concave and the constant C multiplies a strictly convex function, we must have $C \leq 0$. Therefore, p_N^* cannot be smaller than the belief at which the linear part of (5) equals s. This establishes $p_N^* \geq \bar{p}_N$.

To prove the strict inequality, suppose that $p_N^* = \bar{p}_N$ (and hence $C = 0$). Then U_N^* coincides on $[j^{-1}(p_N^*), p_N^*]$ with the expected cost function associated with the (non-Markovian) strategy of having all players switch to the safe arm upon the first breakdown. As these costs tend to $\frac{N\lambda_0}{r+N\lambda_0}\left(\frac{rh}{N} + s\right) > \lambda_0 h$ as p tends to zero, this strategy is strictly suboptimal for small p. As such small p are reached with positive probability under this strategy when we start from a belief in $[j^{-1}(p_N^*), p_N^*]$, U_N^* must be strictly smaller than the linear part of (5) on this interval – a contradiction. ■

The intuition for the lack of smooth pasting at the socially optimal cut-off is exactly the same as in the case of conclusive breakdowns.

4.2 Symmetric Equilibrium

For conclusive breakdowns, we constructed the unique symmetric MPE by pasting together the candidate value functions corresponding to all players playing risky, using an interior allocation, and playing safe, respectively. We did so in the manner of backward induction, moving from lower to higher beliefs p. With inconclusive breakdowns, this is infeasible because the post-jump value $u(j(p))$ is no longer fixed at $\lambda_1 h$ (the expected total cost of a known bad machine) but must itself be determined in equilibrium.

We therefore adopt an alternative approach. Given a belief at which experimentation takes off, and taking the safe cost level as the initial condition to the right of this belief, we solve the relevant differential-difference equation moving from higher to lower beliefs. The equilibrium value function is then pinned down uniquely by the requirement that it lie everywhere in between the single-agent and the cooperative value functions, and hence tend to $\lambda_0 h$ as p goes to zero.[15]

Proposition 5 (Symmetric MPE, $\lambda_0 > 0$) *If breakdowns are inconclusive, the N-player experimentation game has a unique symmetric Markov perfect equilibrium with the common posterior belief as the state variable. The equilibrium strategy is continuous and non-increasing, and there are threshold beliefs $p_N^\dagger > \tilde{p}_N$ with $p_1^* < \tilde{p}_N < p_N^*$ such that all players play risky for $p \leq p_N^\dagger$, use an interior allocation for $p_N^\dagger < p < \tilde{p}_N$, and play safe for $p \geq \tilde{p}_N$. The equilibrium value function is continuous, strictly increasing on $[0, \tilde{p}_N]$, and once continuously differentiable except for a concave kink at $\tilde{p}_N$.*

Proof: For any $\tilde{p} \in [p_1^*, p_N^*]$, let $u_{\tilde{p}} :]0, 1] \to I\!R$ be the unique solution of the differential-difference equation

$$b(p, u) = \max\left\{\frac{\lambda(p)h - u(p)}{N}, c(p)\right\} \tag{6}$$

subject to $u_{\tilde{p}} = s$ on $[\tilde{p}, 1]$.

[15] This 'shooting' method is the same as in Keller and Rady (2010) except for the complication that we are trying to 'hit' a point where the relevant differential equation is singular (the coefficient of the first derivative vanishes at $p = 0$). We overcome it by first constructing a sequence of solutions on subintervals that get closer and closer to $p = 0$, and then showing existence of a convergent subsequence. Earlier examples of this approach can be found in Keller and Rady (1999, 2003) and Bonatti (2011).

We first show that $u_{p_1^*} > U_1^*$ on $]0, p_1^*[$. Noting that $\Delta\lambda\, p_1^*(1-p_1^*)(u_{p_1^*})'(p_1^*)/r = b(p_1^*, u_{p_1^*}) = c(p_1^*) = b(p_1^*, U_1^*) = \Delta\lambda\, p_1^*(1-p_1^*)(U_1^*)'(p_1^*)/r$, we see that $(u_{p_1^*})'(p_1^*) = (U_1^*)'(p_1^*)$. Immediately to the left of p_1^*, moreover,

$$\Delta\lambda\, p(1-p)(u_{p_1^*})'(p) - \lambda(p)\,[s - u_{p_1^*}(p)] = r\,[\lambda(p)h - s]$$

and

$$\Delta\lambda\, p(1-p)(U_1^*)'(p) - \lambda(p)\,[s - U_1^*(p)] = r\,[\lambda(p)h - U_1^*(p)],$$

so that the difference $d = u_{p_1^*} - U_1^*$ solves

$$\Delta\lambda\, p(1-p)d'(p) + \lambda(p)d(p) = r\,[U_1^*(p) - s].$$

Differentiating both sides with respect to p and using the fact that $d(p_1^*) = d'(p_1^*) = 0$ as well as $(U_1^*)'(p_1^*) > 0$, we see that $d''(p_1^*) > 0$, and hence $d > 0$ immediately to the left of p_1^*. Now, suppose that there is a belief in $]0, p_1^*[$ at which $d \leq 0$. Then there exist $p' < p''$ in this interval such that $d(p') = 0$, $d > 0$ on $]p', p_1^*[$, and the restriction of d to $[p', 1]$ assumes a positive global maximum at p''. As $d'(p'') = 0$ and $d(j(p'')) \leq d(p'')$, we have $b(p'', u_{p_1^*}) \geq b(p'', U_1^*)$; as $b(p'', U_1^*) > c(p'')$, moreover, (6) implies $u_{p_1^*}(p'') = \lambda(p'')h - Nb(p'', u_{p_1^*}) \leq \lambda(p'')h - b(p'', U_1^*) = U_1^*(p'')$ – a contradiction.

Analogous steps establish that $u_{p_N^*} < U_N^*$ on $]0, p_N^*[$. By continuous dependence of $u_{\tilde{p}}$ on $\tilde{p}$, we can now find beliefs $\tilde{p}_\ell \in\,]p_1^*, p_N^*[$ for $\ell = 1, 2, \ldots$ such that $u_\ell = u_{\tilde{p}_\ell}$ satisfies $U_N^*(\ell^{-1}) \leq u_\ell(\ell^{-1}) \leq U_1^*(\ell^{-1})$. By the same argument as above, in fact, we have $U_N^* \leq u_\ell \leq U_1^*$ on $[\ell^{-1}, 1]$. Selecting a subsequence if necessary, we can assume that the beliefs $\tilde{p}_\ell$ converge monotonically to some limit $\tilde{p}_\infty$. If $\tilde{p}_\ell$ decreases with ℓ, we set $I_\ell =]\ell^{-1}, p_\infty[$; otherwise we set $I_\ell =]\ell^{-1}, p_\ell[$. In either case, $I_{\ell+1} \supset I_\ell$ for all ℓ, and $\bigcup_{\ell=1}^\infty I_\ell =]0, p_\infty[$.

Next, we note that for each ℓ, there exists a constant $C_\ell > 0$ such that the following holds for any function $u\colon]0, 1] \to I\!\!R$ satisfying $U_N^* \leq u \leq U_1^*$: if u solves (6) on I_ℓ, then $|u'| \leq C_\ell$ on this interval. In fact, the stated conditions imply both

$$\Delta\lambda\, p(1-p)u'(p) \leq \lambda(p)\,[U_1^*(j(p)) - U_N^*(p)] + r \max\left\{\tfrac{\lambda(p)h - U_N^*(p)}{N}, c(p)\right\}$$

and

$$\Delta\lambda\, p(1-p)u'(p) \geq \lambda(p)\,[U_N^*(j(p)) - U_1^*(p)] + r \max\left\{\tfrac{\lambda(p)h - U_1^*(p)}{N}, c(p)\right\}$$

in I_ℓ, from which the claim follows immediately.

This in turn implies that for any $L = 1, 2, \ldots$, the sequences $\{u_\ell\}_{\ell \geq L}$ and $\{u'_\ell\}_{\ell \geq L}$ are uniformly bounded and equicontinuous on I_L. Repeatedly applying the Arzela-Ascoli theorem and then selecting the 'diagonal' subsequence, we obtain a sequence of functions $\{u_L\}_{L=1}^\infty$ and a limit function $\tilde{u}$ such that u_L converges pointwise to $\tilde{u}$ on $]0, 1]$ and u'_L

converges uniformly on each closed subinterval of $]0, \tilde{p}_\infty[$. On the latter interval, therefore, $\tilde{u}$ is once continuously differentiable and solves (6); on $[\tilde{p}_\infty, 1]$, we obviously have $\tilde{u} = s$. As $U_N^* \leq \tilde{u} \leq U_1^*$ on $]0, 1]$, finally, we can extend $\tilde{u}$ continuously to the closed unit interval by setting $\tilde{u}(0) = \lambda_0 h$.

From now, on we write $\tilde{p}_N$ instead of $\tilde{p}_\infty$. In view of what was shown at the start of this proof, we have $p_1^* < \tilde{p}_N < p_N^*$. Letting p tend to $\tilde{p}_N$ from below in (6), we see that $\tilde{u}'(\tilde{p}_N-)$ has the same sign as $c(\tilde{p}_N)$, which is positive since $\tilde{p}_N > p_1^* > p^m$.

We wish to establish that $\tilde{u}$ is strictly increasing on $[0, \tilde{p}_N]$. Suppose that this is not the case. Then there exist beliefs $p' > q'$ in $]0, \tilde{p}_N[$ such that $\tilde{u}(p') - \tilde{u}(q') \leq 0$ is the minimum of $\tilde{u}(p) - \tilde{u}(q)$ on $\{(p, q) \in [0, 1]^2 \colon p \geq q\}$. As $\tilde{u}'(p') = 0$, (6) yields

$$\lambda(p')\left[\tilde{u}(j(p')) - \tilde{u}(p')\right] + r \max\left\{\tfrac{\lambda(p')h - \tilde{u}(p')}{N}, c(p')\right\} = 0.$$

As $\tilde{u}(p') \leq U_1^*(p') < \lambda(p')h$, this implies $\tilde{u}(j(p')) < \tilde{u}(p')$ and hence $\tilde{u}(j(p')) - \tilde{u}(q') < \tilde{u}(p') - \tilde{u}(q')$. As $j(p') > p' > q'$, this is a contradiction.

Now, let $p_N^\dagger = \inf\{p : \tilde{u}(p) > s - (N-1)\, c(p)\}$ and set $\tilde{k}(p) = 1$ for $p \leq p_N^\dagger$, $\tilde{k}(p) = \frac{s - u(p)}{(N-1)\, c(p)}$ for $p_N^\dagger < p < \tilde{p}_N$, and $\tilde{k}(p) = 0$ for $p \geq \tilde{p}_N$. This strategy is non-increasing and Lipschitz-continuous. It is straightforward to verify that all players using this strategy constitutes a symmetric equilibrium with value function $\tilde{u}$.

To establish uniqueness of the symmetric MPE, it is useful to note that the players' common value function in any such equilibrium must solve the variational inequality

$$\max\left\{b(p, u) - \max\left\{\tfrac{\lambda(p)h - u(p)}{N}, c(p)\right\},\ u(p) - s\right\} = 0. \tag{7}$$

To see that this is the case, suppose that all players using a strategy k constitutes a symmetric MPE with equilibrium value function u, and consider the three cases that are possible according to the characterization of best responses in Section 2. First, if $k(p) = 0$ and $u(p) = s$, then (7) holds if and only if $b(p, u) \leq \max\left\{\frac{c(p)}{N}, c(p)\right\}$; as $b(p, u) \geq 0$, this inequality is tantamount to $b(p, u) \leq c(p)$, which must hold because otherwise $k(p) = 0$ would not be a best response against the other $N - 1$ players using the safe arm exclusively. Second, if $0 < k(p) < 1$ and u satisfies $b(p, u) = c(p)$ with $s - (N-1)\, c(p) < u(p) < s$, then (7) holds because $\max\left\{\frac{\lambda(p)h - \tilde{u}(p)}{N}, c(p)\right\} = c(p)$. Third, if $k(p) = 1$ and $u(p) = s - Nb(p, u) + c(p) \leq s - (N-1)\, c(p)$, then (7) holds because $b(p, u) \geq c(p)$ and $\max\left\{\frac{\lambda(p)h - u(p)}{N}, c(p)\right\} = \max\{b(p, u), c(p)\} = b(p, u)$.

Clearly, $\tilde{u}$ solves (7). Suppose that u is also a solution with $u(0) = \lambda_0 h$ and $u(1) = s$. Then a straightforward extension of the arguments given at the start of the proof shows that $u - \tilde{u}$ assumes neither a positive maximum nor a negative minimum. So we must have $u = \tilde{u}$ as claimed. ■

The explanations for the kink of the equilibrium value function at $\tilde{p}_N$, for smooth pasting at $p_N^\dagger$ and for the encouragement effect (that is, $\tilde{p}_N > p_1^*$) are the same as in the case of conclusive breakdowns.

Immediately to the left of $\tilde{p}_N$, (6) has the general solution

$$w(p) = s + \frac{r}{\lambda_1}(\lambda_1 h - s)p - \frac{r}{\lambda_0}(s - \lambda_0 h)(1-p) + C\,(1-p)\left(\frac{1-p}{p}\right)^{\lambda_0/\Delta\lambda},$$

which is strictly convex if and only if $C > 0$. Under the value-matching condition $w(\tilde{p}_N) = s$, this is equivalent to

$$\tilde{p}_N < \frac{\lambda_1}{\Delta\lambda}\,\frac{s - \lambda_0 h}{s}\;.$$

As the right-hand side of this inequality tends to 1 as $\lambda_0 \to 0$, we see that the equilibrium value function is strictly convex immediately to the left of $\tilde{p}_N$ at least for small λ_0. As in the case of conclusive breakdowns, therefore, a non-negative value of information 'in the large' can co-exist with a negative value of information 'in the small'.

4.3 Asymmetric Equilibria

The construction of asymmetric Markov perfect equilibria for inconclusive breakdowns is considerably more difficult than for conclusive ones because we can no longer use backward induction from the belief at which all players start using the risky arm. Moreover, asymmetric actions (and hence asymmetric total expected costs) on some interval of beliefs I necessarily imply asymmetric post-jump continuation values on the interval of more optimistic beliefs $j^{-1}(I)$; and as more optimism translates into a higher intensity of experimentation, the latter interval will be reached with positive probability.

In the scenario with inconclusive good news, by contrast, the beliefs in $j^{-1}(I)$ are more *pessimistic* than those in I; if the interval I on which players take asymmetric actions is close to the belief at which all experimentation stops in equilibrium, the interval $j^{-1}(I)$ is never reached. This allows Keller and Rady (2010) to construct asymmetric equilibria for an arbitrary number of players in which actions and total payoffs are symmetric everywhere except on some interval of beliefs where the players take turns playing risky; see their Proposition 5.

On the path of play in these equilibria, the players always have symmetric continuation values after a breakthrough. If the last experimenter is instead rewarded with a higher payoff after a breakthrough, equilibrium experimentation can be sustained on a larger range of beliefs, as Keller and Rady (2010) illustrate for λ_0 close to zero by means of numerically computed two-player equilibria in simple strategies; see their Section 7. The method used to construct these equilibria carries over to the present framework with two modifications: on the one hand, there is no need to require that λ_0 be close to zero; on the other hand, the construction is more involved in that we again have to solve for the cost functions by 'shooting' into the singularity at $p = 0$, which can be done as for the symmetric MPE.

The details of this construction are available upon request. We do not present them here because it is clear from our earlier arguments that these asymmetric equilibria again exhibit a failure of smooth pasting at the belief where the first experimenter starts using the risky arm.

5 Concluding Remarks

The aim of this paper was to identify and explain the differences between the bad-news and good-news versions of strategic experimentation with Poisson bandits. Consequently, we did not address results that are common to both versions and can be proved exactly as in Keller, Rady and Cripps (2005) and Keller and Rady (2010). These include the non-existence of equilibria in cut-off strategies and (for inconclusive news) the representation of equilibrium value functions in a recursive closed form.

We maintained the assumption made in these earlier papers that the size of a lump-sum payoff or cost conveys no information about the state of the world. If this size were informative, it would no longer be exogenously predetermined whether a news event makes the players more optimistic or more pessimistic. In such a model, it would be impossible, therefore, to construct equilibrium payoff functions iteratively by moving against the direction of jumps in beliefs, as we did in Section 4. The variational inequality (7) that we used to prove uniqueness of the symmetric equilibrium would still hold, however, and could provide a starting point for an existence and uniqueness proof based on the theory of viscosity solutions, for example. We intend to explore this in future work.

References

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

Acemoglu, D., K. Bimpikis and A. Ozdaglar (2011): "Experimentation, Patents, and Innovation," *American Economic Journal: Microeconomics*, **3**, 37–77.

Alili, L. and A.E. Kyprianou (2005): "Some Remarks on First Passage of Lévy Processes, the American Put and Pasting Principles," *Annals of Applied Probability*, **15**, 2062–2080.

Baurdoux, E.J., A.E. Kyprianou and J.C. Pardo (2011): "The Gapeev-Kühn Stochastic Game Driven by a Spectrally Positive Lévy Process," *Stochastic Processes and their Applications*, **121**, 1266–1289.

Bayraktar, E. and H. Xing (2012): "Regularity of the Optimal Stopping Problem for Jump Diffusions," *SIAM Journal of Control and Optimization*, **50**, 1337–1357.

Bergemann, D. and U. Hege (1998): "Venture Capital Financing, Moral Hazard and Learning," *Journal of Banking and Finance*, **22**, 703–735.

Bergemann, D. and U. Hege (2005): "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, **36**, 719–752.

Besanko, D. and J. Wu (2012): "The Impact of Market Structure and Learning on the Tradeoff between R&D Competition and Cooperation," to appear in *Journal of Industrial Economics*.

Biais, B., T. Mariotti, J.-C. Rochet and S. Villeneuve (2010): "Large Risks, Limited Liability, and Dynamic Moral Hazard," *Econometrica*, **78**, 73–118.

Board, S. and M. Meyer-ter-Vehn (2012): "Reputation for Quality," working paper, University of California at Los Angeles.

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

Bolton, P. and C. Harris (2000): "Strategic Experimentation: the Undiscounted Case," in *Incentives, Organizations and Public Economics – Papers in Honour of Sir James Mirrlees*, ed. by P.J. Hammond and G.D. Myles. Oxford: Oxford University Press.53–68

Bonatti, A. (2011): "Menu Pricing and Learning," *American Economic Journal: Microeconomics*, **3**, 124–163.

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

Boyarchenko, S. and S. Levendorskiĭ (2002): "Perpetual American Options under Lévy Processes," *SIAM Journal of Control and Optimization*, **40**, 1663–1696.

Cohen, A. and E. Solan (2012): "Bandit Problems with Lévy Payoff Processes," to appear in *Mathematics of Operations Research*.

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

Dixit, A.K. (1993): *The Art of Smooth Pasting.* Harwood: Academic Press.

Dixit, A.K. and R.S. Pindyck (1994): *Investment under Uncertainty.* Princeton: Princeton University Press.

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

Dumas, B. (1991): "Super Contact and Related Optimality Conditions," *Journal of Economic Dynamics and Control*, **15**, 675–695.

Faingold, E. and Y. Sannikov (2011): "Reputation in Continuous-Time Games," *Econometrica*, **79**, 773–876.

Fudenberg, D. and D.K. Levine (2007): "Continuous Time Limits of Repeated Games with Imperfect Public Monitoring," *Review of Economic Dynamics*, **10**, 173-192.

Gapeev, P.V. and C. Kühn (2005): "Perpetual Convertible Bonds in Jump-Diffusion Models," *Statistics and Decisions*, **23**, 15-31.

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

Hopenhayn, H.A. and F. Squintani (2011): "Preemption Games with Private Information," *Review of Economic Studies*, **78**, 667–692.

Keller, G. and S. Rady (1999): "Optimal Experimentation in a Changing Environment," *Review of Economic Studies*, **66**, 475–507.

Keller, G. and S. Rady (2003): "Price Dispersion and Learning in a Dynamic Differentiated-Goods Duopoly," *RAND Journal of Economics*, **34**, 138–165.

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

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

Klein, N. and S. Rady (2011): "Negatively Correlated Bandits," *Review of Economic Studies*, **78**, 693–732.

Ludkovski, M. and R. Sircar (2012): "Exploration and Exhaustibility in Dynamic Cournot Games," *European Journal of Applied Mathematics*, **23**, 343–372.

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

Mikhalevich, V.S. (1958): "A Bayes Test of Two Hypotheses Concerning the Mean of a Normal Process," *Visnik Kiivskogo Nacionalnogo Universitetu*, **1**, 254–264 (in Ukrainian).

Moscarini, G. and F. Squintani (2010): "Competitive Experimentation with Private Information: The Survivor's Curse," *Journal of Economic Theory*, **145**, 639–660.

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

Peskir, G. and A.N. Shiryaev (2000): "Sequential Testing Problems for Poisson Processes," *Annals of Statistics*, **28**, 837–859.

Peskir, G. and A. Shiryaev (2006): *Optimal Stopping and Free-Boundary Problems*. Basel: Birkhäuser Verlag.

Rosenberg, D., E. Solan and N. Vieille (2007): "Social Learning in One-Armed Bandit Problems," *Econometrica*, **75**, 1591–1611.

Samuelson, P.A. (1965): "Rational Warrant Pricing," *Industrial Management Review*, **6**, 41–50.

Strulovici, B. (2010): "Learning While Voting: Determinants of Collective Experimentation," *Econometrica*, **78**, 933–971.

Strulovici, B. and M. Szydlowski (2012): "On the Smoothness of Value Functions and the Existence of Optimal Strategies," working paper, Northwestern University.