Rochester Center for Economic Research

Keeping Your Options Open

Forand, Jean Guillaume

Working Paper No. 557
October 2010

UNIVERSITY OF
ROCHESTER

Keeping Your Options Open*

Jean Guillaume Forand†

October 16, 2010

Abstract

In standard models of experimentation, the costs of project development consist of (*i*) the direct cost of running trials as well as (*ii*) the implicit opportunity cost of leaving alternative projects idle. Another natural type of experimentation cost, the cost of holding on to the option of developing a currently inactive project, has not been studied. In a (multi-armed bandit) model of experimentation in which inactive projects have explicit maintenance costs and can be irreversibly discarded, I fully characterise the optimal experimentation policy and show that the decision-maker's incentive to actively manage its options has important implications for the order of project development. In the model, an experimenter searches for a success among a number of projects by choosing both those to develop now and those to maintain for (potential) future development. In the absence of maintenance costs, the optimal experimentation policy has a 'stay-with-the-winner' property: the projects that are more likely to succeed are developed first. Maintenance costs provide incentives to bring the option value of less promising projects forward, and under the optimal experimentation policy, projects that are less likely to succeed are sometimes developed first. A project development strategy of 'going-with-the-loser' strikes a balance between the cost of discarding possibly valuable options and the cost of leaving them open.

1 Introduction

When experimentation is costly, decision-makers must choose which alternatives to actively investigate and which to leave 'on the back burner'. Consider, for example, a firm engaged in research and development facing many technologies that can lead to comparable innovations. Investing in multiple technologies simultaneously is costly, so the firm prioritises its allocation of funds to competing ideas. A massive number of books and business articles on project management help managers decide which technologies to develop and, more importantly, whether/when to transfer resources to other projects following disappointing results in priority projects. As another

*A previous draft was circulated under the title 'Experimentation with Costly Project Maintenance'. I would like to thank Li Hao, Martin Osborne and Colin Stewart for their supervision, comments and suggestions. I would also like to thank Ettore Damiano, John Duggan, Carolyn Pitchik, Wing Suen and seminar participants at the University of Waterloo and the Fall 2010 METC.

†W. Allen Wallis Institute, 107 Harkness Hall, Box 027-0158, University of Rochester, NY 14627-0158. jg-forand@yahoo.ca.

example, professional sports teams' in-game roster management decisions stem from exogenous restrictions on their ability to learn about multiple players at once. A single player can play at a given position at any given time during a game and coaches/managers can gather information about their players' abilities only by having them replace a teammate.

In standard models of experimentation, the choice of gathering information about one alternative as opposed to another entails only an implicit opportunity cost: the foregone opportunity of learning about the inactive alternative. However, retaining the option to investigate a currently shelved alternative often involves explicit maintenance costs. Firms engaged in research and development routinely devote resources solely to keep open the option of developing a technology that is currently 'on the back burner', which involves the costly upkeep of specialised equipment and paying the salaries of skilled workers or scientists that can be lost to other firms. In professional sports, the option to develop players of unknown quality is kept open by filling roster spots with 'bench' players, who may seldom get the opportunity to play but command millions of dollars' worth of salaries.

In this paper, I present a simple (multi-armed bandit) model of experimentation in which projects (arms) that are not being developed (pulled) have explicit maintenance costs. The experimenter is thus led to actively manage its set of options, as it faces a choice between paying to keep some options open or discarding them (irreversibly) altogether. Discarding an inactive project liquidates its option value, which is realised in the event that currently active projects are deemed unpromising. To avoid both destroying this option value and paying to maintain it, the experimenter has an incentive to bring it forward by altering the order of project development.

In a tractable setup in which two risky projects can be either good or bad and only good projects eventually succeed if developed, I fully characterise the optimal experimentation policy with maintenance costs and show that it entails significant departures from standard results. In the absence of maintenance costs, the optimal experimentation policy has the well-known 'stay-with-the-winner' property: the project that is more likely to succeed is investigated first.[1] In the presence of maintenance costs, 'going-with-the-loser' can be optimal: projects less likely to succeed may be investigated first. When 'going-with-the-loser', the experimenter brings the option value of 'losing' projects forward through a simple culling rule. Such projects are granted a 'last chance' to succeed through a short and intense period of experimentation, after which they are permanently discarded in favour of more promising projects.

While the idea of maintenance costs is natural and widely applicable, one way to interpret these results is as providing some rational foundations for the behaviour described as 'throwing good money after bad' or escalation, in which decision-makers fail to 'know when to pull the plug' and appear to cling to projects that have repeatedly failed to achieve results.[2] Common explanations have revolved around decision-makers falling prey to some form of sunk cost fal-

[1]The term 'stay-with-the-winner' is coined by Berry and Fristedt (1985).
[2]See Staw (1981), Staw and Ross (1987) and Garland (1990).

lacy. However, it need not be the case that observing an intensified commitment to a given project following a failure is the result only of non-rational behaviour.[3] Indeed, in my model, the experimenter throws good money after bad precisely in order to convince itself that the initial investments were indeed a bad idea, thus ensuring a quicker extrication of resources from a hopeless project towards more promising ones.

The following example illustrates the main lessons of the paper by clarifying why experimenting with the 'losing' project first can be optimal when maintaining inactive projects is costly. An experimenter can devote a trial to one of two projects, A and B, in each of three periods. Projects are risky in that the payoffs they deliver are unknown. A project of type *Good* delivers a one-time payoff of 1 with probability $G > 0$ in any trial. Assume that experimentation ends once a single trial is successful. Direct experimentation costs are $\overline{k} > 0$ per trial, maintenance costs for an unused project are $\underline{k} \geq 0$ per period and there is no discounting.

A project's current state is characterised by the experimenter's belief that it is of type *Good* and repeated failures make the experimenter more pessimistic about the project. Let p_J^i be the probability that project J is of type *Good* given that it has failed i trials, with $i \in \{0, 1, 2\}$. By Bayes' rule, $p_J^i = \frac{p^{i-1}(1-G)}{1-p^{i-1}G}$ for $i = \{1, 2\}$, and initial beliefs (p_A^0, p_B^0) are given. Assume that $p_B^0 \in [p_A^2, p_A^1]$. This ensures that project A is the better project ex ante and that in the absence of maintenance costs, that is if $\underline{k} = 0$, the optimal experimentation sequence must develop project A twice and project B once. While all such experimentation sequences generate the same probability of a success, the 'go-with-the-winner' sequence AAB yields the highest payoff as it maximises the probability that a success arrives early and avoids further experimentation costs.

When $\underline{k} > 0$, it is straightforward to show that the optimal experimentation sequence will always be one of $AA|_AB$, $|_BAAA$ or $B|_BAA$, where $|_J$ represents the discarding of project J. That is, either the experimenter sticks with the 'go-with-the-winner' rule, abandons the 'losing' project B immediately or it gives project B an early chance to succeed and discards it following a failure. Let $V(s; p_A^0, p_B^0)$ be the expected payoff to experimentation sequence s given initial beliefs (p_A^0, p_B^0). Then

$$\begin{aligned}
V(AA|_AB; p_A^0, p_B^0) &= p_A^0 + (1-p_A^0)p_A^1 + (1-p_A^0)(1-p_A^1)p_B^0 \\
&\qquad - \Big[(\overline{k}+\underline{k}) + (1-p_A^0)(\overline{k}+\underline{k}) + (1-p_A^0)(1-p_A^1)\overline{k}\Big], \\
V(B|_BAA; p_A^0, p_B^0) &= p_B^0 + (1-p_B^0)p_A^0 + (1-p_B^0)(1-p_A^0)p_A^1 \\
&\qquad - \Big[(\overline{k}+\underline{k}) + (1-p_B^0)\overline{k} + (1-p_B^0)(1-p_A^0)\overline{k}\Big], \\
V(|_BAAA; p_A^0, p_B^0) &= p_A^0 + (1-p_A^0)p_A^1 + (1-p_A^0)(1-p_A^1)p_A^2 \\
&\qquad - \Big[\overline{k} + (1-p_A^0)\overline{k} + (1-p_A^0)(1-p_A^1)\overline{k}\Big].
\end{aligned}$$

[3] In the management literature, Bowen (1987) has related misgivings about such interpretations of escalation behaviour. See also McAfee et al. (2010).

The incentive to bring the option value of project B forward entails a trade-off. The benefits of experimentation sequence $B|_B AA$ are that (i) (relative to sequence $|_B AAA$) the value of project B gets exploited and the decision to discard B is better informed while (ii) (relative to sequence $AA|_A B$) saving on maintenance costs. However, to the experimentation sequence $B|_B AA$ are associated both (i) the maintenance cost (relative to $|_B AAA$) and (ii) the opportunity cost (relative to $AA|_A B$) of leaving the 'better' project A idle while experimenting with project B.

Simple calculations show that $V(AA|_A B; p_A^0, p_B^0) - V(B|_B AA; p_A^0, p_B^0)$ is decreasing in p_B^0. Hence, if $B|_B AA$ is preferred to $AA|_A B$ for some $p_B^0 \in [p_A^1, p_A^2]$, then this is also the case for all $p_B'^0 > p_B^0$. Note also that $V(|_B AAA; p_A^0, p_B^0)$ is independent of p_B^0 and that $V(|_B AAA; p_A^0, p_B^0) - \max\{V(AA|_A B; p_A^0, p_B^0), V(B|_B AA; p_A^0, p_B^0)\}$ is decreasing in p_B^0 and is strictly positive at $p_B^0 = p_A^1$. That is, when $p_B^0 = p_A^1$, all three experimentation sequences $AA|_A B$, $B|_B AA$ and $|_B AAA$ yield the same success probabilities, yet $|_B AAA$ has strictly lower costs. Hence, for fixed G, $\overline{k}$, $\underline{k}$ and p_A^0, the optimal experimentation policy can be represented by beliefs $\underline{p}$, $\overline{p}$ with $p_A^2 \leq \underline{p} \leq \overline{p} \leq p_A^1$, such that $|_B AAA$ is optimal on $[p_A^2, \underline{p}]$, $AA|_A B$ is optimal on $[\underline{p}, \overline{p}]$ and $B|_B AA$ is optimal on $[\overline{p}, p_A^1]$. In general, all three intervals can be non-empty. An example has $G = \frac{2}{5}$, $\overline{k} = \frac{6}{100}$, $\underline{k} = \frac{3}{200}$ and $p_A^0 = \frac{45}{100}$. Then it can be computed that $p_A^1 \approx 0.33$ and $p_A^2 \approx .23$, while $\underline{p} \approx 0.32$ and $\overline{p} \approx 0.29$. This is depicted in Figure 1.

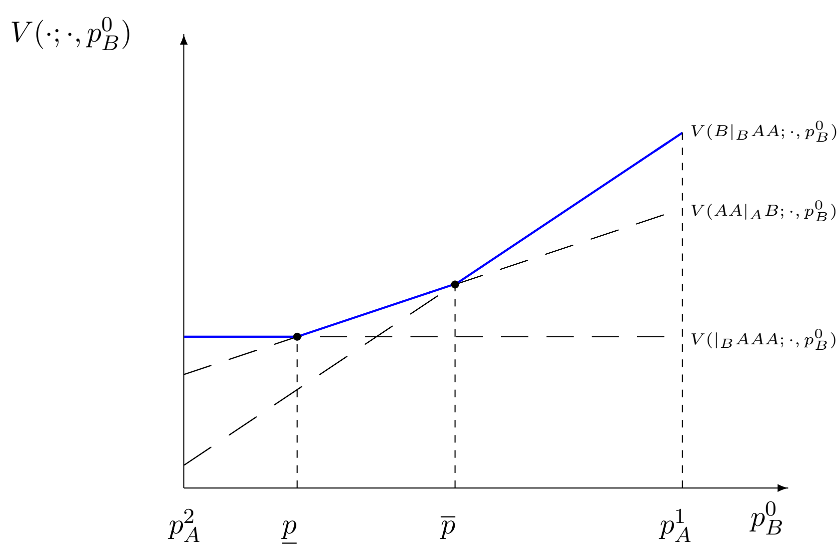

Figure 1: Example: Optimal Experimentation as a function of p_B^0.

I model experimentation as a multi-armed bandit problem.[4] In the standard discounted

[4]See Berry and Fristedt (1985). Bergemann and Välimäki (2006) survey the bandits literature with an eye to applications in economics.

bandit problem with independent arms, the optimal experimentation policy is the well-known (Gittins) index policy,[5] which is not robust to perturbations of the model such as correlated arms, non-geometric discounting and the simultaneous pulling of multiple arms. Much less is known about optimal experimentation policies when these do not take the simple index form. Closer to this paper, Banks and Sundaram (1994) have shown that index policies are not optimal in the presence of switching costs between arms.[6] Switching costs are attributed to an inactive arm only when experimentation transitions to it and are always accompanied by an observation from that arm. Maintenance costs, on the other hand, need to be paid whenever an inactive arm is not pulled and never generate observations from that arm. Nevertheless, the bandit problem with maintenance costs fails to admit a Gittins index representation for the reason found by Banks and Sundaram (1994): the index of a given maintained arm would have to be a function of the maintenance cost, and this relationship would depend nontrivially on the characteristics of outside arms.[7]

I adopt the exponential bandit framework due to Keller et al. (2005), which yields a continuous time infinite horizon version of the model in the example from above. Exponential bandits have proved useful in applications due to their tractability. Keller et al. (2005), following Bolton and Harris (1999), study strategic experimentation and the free-riding incentives of multiple agents facing a single risky arm. Keller and Rady (2009) generalise the model to 'poisson' bandits that allow for arms of the bad type to also generate successes. Klein and Rady (2008) allow for each of two experimenters to have perfectly negatively correlated versions of the same risky arm. Strulovici (2009) applies the model in a voting framework. Bergemann and Hege (1998) introduce a discrete-time version of the model to study the moral hazard problem arising between bankers (principal) and venture capitalists (experimenters). In this vein, recent papers by Bonatti and Hörner (2009) and Hörner and Samuelson (2009) focus on the provision of incentives to experimenting agents. Bonatti and Hörner (2009) derive another version of the 'stay-with-the-winner' rule when agents can experiment with multiple disjunctive projects, i.e., when only a single project success is required. They also uncover a 'go-with-the-loser' rule when projects are conjunctive, i.e., when success on both projects is required. In that case, experimenting first with the losing project is optimal since a success on the winning project is worthless on its own. My results show that with maintenance costs to inactive projects, 'going-with-the-loser' is optimal even with disjunctive projects.

In Section 2, I describe the model. In Section 2.1, I extend the standard expressions for

[5] To each arm is assigned a number (index) that depends only on the ex ante characteristics and accumulated observations of that project. The optimal experimentation policy consists of always selecting a project among those with maximal indices.

[6] General characterisations of optimal experimentation policies with switching costs have proven difficult to obtain. For details, see Jun (2004). An exception is Bergemann and Välimäki (2001), who exploit results of Banks and Sundaram (1992b) on bandits with a countable numbers of ex ante identical arms to show that an experimenter never switches back to an arm it switched away from earlier.

[7] It is not clear how to define an index policy in the presence of maintenance costs since experimentation policies need to specify both which arm is pulled and which arms are maintained.

the experimenter's optimal payoffs in the continuous-time exponential two-armed (optimal stopping) bandit problem to the case with two risky projects. In Section 3, I characterise optimal experimentation in a benchmark model in which inactive projects have maintenance costs but the experimenter cannot discard risky projects individually. This is equivalent to a standard three-armed bandit with two risky arms and no maintenance costs, and I show that the optimal (Gittins index) experimentation policy involves the 'stay-with-the-winner' rule. That is, conditional on continued experimentation, it is optimal to select the project most likely to succeed. In Section 4, I present the main results of the paper for the model in which inactive projects have maintenance costs and the experimenter can discard individual projects. First, I show that if the optimal policy ever 'goes-with-the-loser', it will do so in a very specific way, notably through a culling rule. The losing project will be chosen continuously for a short period, after which, in the absence of a success, it will be discarded. In other words, losing projects are put to trial before winning projects only if they are being granted a 'last chance' to succeed, else, as in the example, they are either maintained but not put to trial or simply discarded. Second, I give a complete characterisation of the optimal policy and show that 'going-with-the-loser' is a robust feature of optimal experimentation with maintenance costs. More precisely, whenever it is not the case that maintenance costs are high enough that the losing project is always discarded immediately, it will be put to trial before the winning arm in non-negligible regions of the belief space.

In Section 5, I show that my results can be extended in two natural directions. First, in the case in which the experimenter has more than two risky projects, the culling rule for losing projects takes a more general form. When the experimenter has three risky projects ranked by their likelihoods of success,[8] I show that if it is ever optimal to experiment with the middle-ranked project, then experimentation can proceed to the top-ranked project only when both the middle-ranked and the lowest-ranked projects have been discarded. In other words, experimenting with a middle-ranked project grants a 'last chance' to all projects of a lower or equal rank. Second, I show that 'going-with-the-loser' is still optimal if successes on various projects are not perfect substitutes but can be accumulated.[9] Hence, my results are due to the incentive to economise on maintenance costs by bringing the option values of inactive projects forward, and not to the fact that a leftover project is rendered valueless by another project's success.

2 Model

Consider a continuous time three-armed bandit problem with two risky arms, A and B, and a safe arm S. Arms will henceforth be referred to as projects. A trial consists of experimenting with a risky project for some time interval $[t, t + dt]$. Trials yield either successes or failures. The type of a risky project is $\theta \in \{Good, Bad\}$. A risky project of type θ that is pulled continuously

[8] The argument is easily extended to more than three projects.

[9] This corresponds to the distinction between disjunctive and additive projects in the language of Bonatti and Hörner (2009).

in time interval $[t, t+dt]$ succeeds with probability Gdt for some $G > 0$ if $\theta = Good$, while it fails for sure if $\theta = Bad$. The types of risky projects A and B are drawn independently. Let $p_J(0)$ be the ex ante probability that project J is of type *Good*. A safe project S yields a flow payoff of 0. A success on either risky project yields a lump-sum payment of 1 and ends the experimentation process.

Experimenting continuously with risky project J in time interval $[t, t+dt]$ entails experimentation cost $\overline{k}dt$. I introduce explicit costs to maintaining inactive risky projects. That is, a risky project that is maintained but not involved in a trial in time interval $[t, t+dt]$ entails a cost of $\underline{k}dt$. The experimenter can irreversibly discard risky projects without cost. That is, it can avoid paying for the maintenance of inactive projects but only at the cost of permanently abandoning some of its options. There are no costs to the safe project, which can be interpreted as an option to quit the experimentation process. The experimenter discounts future payoffs at rate r.

Since experimentation ends after the first success, the only histories after which the experimenter selects a project to experiment with are intervals of time in which only failures have been observed. Strategies should properly be defined on histories. However, any such strategy can be redefined to depend solely on time in the absence of a success. A *strategy* is a collection (α, ϕ_A, ϕ_B) for some function $\alpha : \mathbf{R}_+ \to [0,1] \cup \{S\}$ and decreasing functions $\phi_J : \mathbf{R}_+ \to \{0,1\}$ for $J \in \{A, B\}$. The function α is an *assignment rule* and $\int_t^{t+dt} \alpha(t)$ specifies the fraction of time devoted to experimenting with projects A in time interval $[t, t+dt]$ if the experimenter conducts trials in that interval, while $\alpha(t) = S$ if the experimenter pulls the safe project at time t. The principal is allowed to share the responsibility for the project between the agents in any interval of time. The assumption that the experimenter cannot share the assignment between all three projects and must decide first whether to conduct trials and then how to share experimentation between risky projects is made to simplify the exposition and is in fact without loss of generality for optimal experimentation. Functions ϕ_A and ϕ_B specify *maintenance rules*, with $\phi_J(t) = 1$ if and only if J is maintained at time t. Strategy (α, ϕ_A, ϕ_B) is *admissible* if each component is right-continuous and piecewise Lipschitz continuous. Let $t_J \in [0, \infty) = \sup\{t : \phi_J(t) = 1\}$. Given any initial beliefs $(p_A(0), p_B(0)) \in [0,1]^2$, an admissible strategy (α, ϕ_A, ϕ_B) induces a uniquely defined and continuously differentiable laws of motion for beliefs $(p_A(t), p_B(t))$. These laws of motion, which are given by

$$\dot{p}_A(t) = \begin{cases} -\alpha(t) G p_A(t)(1 - p_A(t)) \text{ for } t \in [0, t_A), \\ 0 \text{ for } t \geq t_A, \end{cases}$$

$$\dot{p}_B(t) = \begin{cases} -(1 - \alpha(t)) G p_B(t)(1 - p_B(t)) \text{ for } t \in [0, t_B), \\ 0 \text{ for } t \geq t_B, \end{cases}$$

are derived in a straightforward way by requiring that the evolution of beliefs be consistent with α and Bayes' rule, and follows Keller et al. (2005).

For much of the paper, it will be more convenient to work with Markov strategies, which are conditioned on the state variable, which is the current beliefs along with the set of maintained projects. Using dynamic programming methods allows for simple expressions for optimal payoffs. However, many of the arguments regarding when and why maintained projects should be discarded are naturally established by considering time paths of play. More formally, a *state* consists of $(p_A, p_B, I_A, I_B) \in [0,1]^2 \times \{0,1\}^2$. A *Markov assignment* is a function $\beta : [0,1]^2 \times \{0,1\}^2 \to [0,1] \cup \{S\}$. *Markov maintenance rules* are functions $\varphi_J : [0,1]^2 \times \{0,1\}^2 \to \{0,1\}$ for $J \in \{A, B\}$ such that $\varphi_J(p_A, p_B, I_A, I_B) = 0$ whenever $I_J = 0$.

Imposing admissibility requirements directly on Markov strategies is cumbersome.[10] A further difficulty in my framework is to determine how the monotonicity (irreversibility) requirements on maintenance rules carry over to restrictions on Markov maintenance rules. To get around these issues, I rely on the admissibility requirement already stated for strategies. Markov strategy $(\beta, \varphi_A, \varphi_B)$ will be said to be *admissible* if given any state (p_A, p_B, I_A, I_B) and initial beliefs $(p_A(0), p_B(0)) = (p_A, p_B)$, there exists a corresponding admissible strategy (α, ϕ_A, ϕ_B) such that for all t

$$
\begin{aligned}
\alpha(t) &= \beta(p_A(t), p_B(t), \phi_A(t), \phi_B(t)),\\
\phi_A(t) &= \varphi_A(p_A(t), p_B(t), \phi_A(t), \phi_B(t)),\\
\phi_B(t) &= \varphi_B(p_A(t), p_B(t), \phi_A(t), \phi_B(t)).
\end{aligned}
$$

Henceforth I will not explicitly restrict the experimenter to using admissible Markov strategies, but I will verify that the optimal Markov strategies I derive, as well as the deviating strategies that support various proofs, are admissible.

A Markov strategy (β, φ) is *symmetric* if

$$
\begin{aligned}
\beta(p_B, p_A, I_B, I_A) &= \begin{cases} 1 - \beta(p_A, p_B, I_A, I_B) & \text{if } \beta(p_A, p_B, I_A, I_B) \neq S, \\ S & \text{if } \beta(p_A, p_B, I_A, I_B) = S, \end{cases}\\
\varphi_J(p_B, p_A, I_B, I_A) &= \varphi_{-J}(p_A, p_B, I_A, I_B) \quad \text{for } J \in \{A, B\}.
\end{aligned}
$$

Given any optimal strategy (β^*, φ^*), there exists an optimal symmetric strategy that achieves the same payoffs. Hence, restricting to symmetric strategies is without loss of generality for the experimenter's payoffs. Given the restriction to symmetric strategies, it is without loss of generality to assume that $p_A \geq p_B$. Henceforth, project A will always be the 'winning' project, with project B the 'losing' project.

Let $W(\alpha, \phi; t, \tau)$ be the experimenter's payoff at time t to strategy (α, ϕ) if a success arrives at time $\tau < \min\{t_A, t_B\}$

$$
W(\alpha, \phi; t, \tau) = e^{-r\tau} - \int_t^\tau e^{-rs}(\overline{k} + \underline{k})ds,
$$

[10] See Fleming and Rishel (1975), Theorem 6.1.

while if a success arrives at $\tau \in [\min\{t_A, t_B\}, \max\{t_A, t_B\}]$

$$W(\alpha, \phi; t, \tau) = e^{-r\tau} - \int_t^{\min\{t_A,t_B\}} e^{-rs}(\overline{k} + \underline{k})ds - \int_{\min\{t_A,t_B\}}^{\tau} e^{-rs}\overline{k}ds,$$

and finally if a success never arrives

$$W(\alpha, \phi; t, \tau) = -\int_t^{\min\{t_A,t_B\}} e^{-rs}(\overline{k} + \underline{k})ds - \int_{\min\{t_A,t_B\}}^{\max\{t_A,t_B\}} e^{-rs}\overline{k}ds.$$

The expected payoff to strategy (α, ϕ) given belief $(p_A(0), p_B(0))$ is

$$V(\alpha, \phi; t) \equiv \mathbf{E}_\tau W(\alpha, \phi; t, \tau),$$

where the expectation is taken over the distribution of stopping times τ determined by (α, ϕ) and $(p_A(s), p_B(s))_t$. Consider an admissible Markov strategy (β, φ) and its corresponding strategy (α, ϕ) for some state (p, I). The expected payoff to (β, φ) in state (p, I) is given by

$$v(\beta, \varphi; p, I) \equiv V(\alpha, \phi; 0)$$

The objective of the experimenter is to find a payoff-maximising strategy. To this end, let $U(t) = \max_{(\alpha,\phi)} V(\alpha, \phi; t)$. Similarly, let $u(p, I) = \max_{(\beta,\varphi)} v(\beta, \varphi; p, I)$.

2.1 Preliminaries: Optimal Payoff Functions

Keller et al. (2005) provide simple expressions for optimal value functions for the two-armed exponential bandit (optimal stopping) problem. In this section, I build on these results to derive the expressions satisfied by the optimal payoff u that will support the characterisations of Sections 3 and 4. To simplify notation, let the number of beliefs listed in a state implicitly denote the set of maintained projects. Hence (p_A, p_B) stands for state $(p_A, p_B, 1, 1)$, (p_A) for state $(p_A, p_B, 1, 0)$ given any p_B, and so on.

The optimal payoff u must satisfy the following Bellman equation

$$u(p_A, p_B) = \max\Bigg\{e^{-rdt}u(p_A, p_B),\ u_A(p_A),\ u_B(p_B),\ \max_{\beta\in[0,1]}\Big\{[\beta p_A G + (1-\beta)p_B G - (\overline{k} + \underline{k})]dt + e^{-rdt}\mathbf{E}[u(p_A + dp_A, p_B + dp_B)|p_A, p_B]\Big\}\Bigg\}. \tag{1}$$

The first term in the brackets of (1) corresponds to the option of employing the safe project in a time interval of length dt. The second and third terms correspond to the options of discarding projects A and B respectively, where u_J corresponds to the optimal payoff to the two-armed bandit problem with risky project J and the safe project. The final term corresponds to the payoff from maintaining both projects and allocating the experimentation effort optimally.

When a risky project has been discarded, the payoff u_J solves

$$u_J(p_J) = \max \left\{ e^{-rdt} u_J(p_J),\ [p_J G - \overline{k}]dt + e^{-rdt}\mathbf{E}[u_J(p_J + dp_J)|p_J] \right\}.$$

The probability of a success in an interval of length dt is $p_J G dt$, and the payoff to a success is 1. The probability of failure is $1 - p_J G dt$. In case of failure, the payoff to the experimenter is $u_J(p_J) + u_J'(p_J)dp_J$, which is equal to $u(p_J) - u_J'(p_J)p_J(1-p_J)Gdt$. By rewriting and cancelling dominated terms

$$ru_J(p_J) = \max \left\{ 0, p_J G - \overline{k} - u_J'(p_J)Gp_J(1-p_J) - u_J(p_J)Gp_J \right\}.$$

Hence, in an open region of beliefs in which project J is used, u_J satisfies the differential equation

$$u_J(p_J)(r + Gp_J) = p_J G - \overline{k} - u_J'(p_J)Gp_J(1-p_J), \tag{2}$$

which can be solved to yield

$$u_J(p_J) = \tilde{C}_J \left(\frac{1-p_J}{p_J} \right)^{\frac{r}{G}} (1-p_J) + p_J \frac{G - \overline{k}}{r+G} - (1-p_J)\frac{\overline{k}}{r}, \tag{3}$$

with the constant of integration $\tilde{C}_J = \left(\frac{\overline{k}}{G-\overline{k}}\right)^{\frac{r}{G}} \frac{G\overline{k}}{r(r+G)}$ and the stopping belief $p_J^* = \frac{\overline{k}}{G}$ determined by value-matching and smooth-pasting conditions

$$u_J(p_J^*) = 0, \text{ and}$$
$$u_J'(p_J^*) = 0.$$

The setup here is slightly different than in Keller et al. (2005), but the expression (3) admits the same interpretation. The term $p_J \frac{G-\overline{k}}{r+G} - (1-p_J)\frac{\overline{k}}{r}$ is the payoff to risky project J in the absence of the ability to quit experimentation, while the term $\tilde{C}_J \left(\frac{1-p_J}{p_J}\right)^{\frac{r}{G}} (1-p_J)$ captures the option value of the quitting option S.

Note that the part of value function (1) in which both projects are maintained is linear in β. Hence, in an open region of the state space in which both projects are maintained, the optimal value is attained for $\beta \in \{0,1\}$, and (1) can be rewritten as

$$ru(p_A, p_B) = \max \left\{ p_A G - (\overline{k} + \underline{k}) - \frac{\partial u(p_A,p_B)}{\partial p_A} Gp_A(1-p_A) - u(p_A,p_B)Gp_A, \right.$$
$$\left. p_B G - (\overline{k} + \underline{k}) - \frac{\partial u(p_A,p_B)}{\partial p_B} Gp_B(1-p_B) - u(p_A,p_B)Gp_B \right\}. \tag{4}$$

Contrary to (2), partial differential equation (4) does not have a simple solution, since such a solution must include an optimal rule for the allocation of trials among projects.

I approach the solution to (4) by abstracting from allocation rules in order to reduce the two-dimensional problem (4) to suitably defined single-dimensional problems. First consider an open region of the state space in which project A is put to trial but both projects are maintained. Then since the optimal Markov strategy is admissible there exists $t' > 0$ and a parametrised path $(p_A(t), p_B)$ such that $U(t) = u(p_A(t), p_B)$ for $t \in (0, t')$. An argument similar to that establishing (2) shows that $U(t)$ satisfies

$$U(t)[r + p_A(t)G] - U'(t) = p_A(t)G - (\overline{k} + \underline{k}). \tag{5}$$

For path $(p_A(t), p_B)$, define $u_A(p_A(t); p_B) \equiv U(t)$. Then $U'(t) = -u'_A(p_A(t); p_B)Gp_A(t)(1 - p_A(t))$, which uses the law of motion for p_A. Condition (5) can be rewritten, eliminating the dependence on time, as

$$u_A(p_A; p_B)[r + p_A G] + u'_A(p_A; p_B)Gp_A(1 - p_A) = p_A G - (\overline{k} + \underline{k}). \tag{6}$$

As for (2), (6) can be solved to yield

$$u_A(p_A; p_B) = C_A(p_B)\left(\frac{1 - p_A}{p_A}\right)^{\frac{r}{G}}(1 - p_A) + p_A\frac{G - (\overline{k} + \underline{k})}{r + G} - (1 - p_A)\frac{\overline{k} + \underline{k}}{r}. \tag{7}$$

In (7), the constant of integration, and hence the option value of project B, will in general depend on p_B, since p_B can affect the payoffs when exiting the A-assignment region. If the parametrised path $(p_A(t), p_B)$ exits the A-assignment region in state (p_A^*, p_B), then p_A^* and $C_A(p_B)$ satisfy the value-matching and smooth-pasting properties

$$\begin{aligned} u_A(p_A^*; C_A(p_B)) &= u(p_A^*, p_B), \text{ and} \\ \frac{\partial}{\partial p_A}u_A(p_A^*; C_A(p_B)) &= \frac{\partial}{\partial p_A}u(p_A^*, p_B). \end{aligned}$$

In general, $u(p_A^*, p_B)$ is endogenous and depends on the experimentation policy once exit from the A-region occurs. If, for example, experimentation exits the A-region into the quitting region at p_A^*, then $u(p_A^*, p_B) = 0$ and $\frac{\partial}{\partial p_A}u(p_A^*, p_B) = 0$, which yields that $p_A^* = \frac{\overline{k}+\underline{k}}{G}$.

Equation (7) establishes a useful necessary condition for optimal payoffs: when project A is put to trial in the optimal solution, payoffs evolve as though the experimenter was facing an optimal stopping problem with cost $\overline{k} + \underline{k}$ for the risky project, with the value to the stopping region adjusted to incorporate continuation payoffs. It will be useful in the sequel to distinguish a payoff of the form (7) from the optimal payoff $u(p_A, p_B)$ in an A-assignment region. Given (p_A, p_B) and some function $C_A(p_B)$, define the righthand side of (7) as $v_A(p_A; C_A(p_B))$.

3 Benchmark: Staying-With-The-Winner

To highlight the impact of maintenance costs on optimal experimentation, a useful benchmark is the case in which risky projects cannot be discarded individually. In this problem, the experimenter can only coarsely manage its options: it can either experiment at the cost of maintaining both projects, or quit experimentation by moving to the safe project and discarding both risky projects. To this end, suppose that the experimenter is restricted to Markov strategies with $\varphi(p, I) = (1, 1)$ for all states (p, I) such that $\beta(p, I) \neq S$. Note that this problem is equivalent to the standard three-armed bandit problem with direct experimentation flow cost $(\overline{k} + \underline{k})$.

Conditional on continuing experimentation, how should the experimenter allocate trials between risky projects? The next lemma shows that the experimenter should always put the project with the highest belief to trial, that is, it should follow a 'stay-with-the-winner' rule. When beliefs are such that $p_A > p_B$, this means using project A. When beliefs are such that $p_A = p_B$, then both projects are 'winning' and 'staying-with-the-winner' entails sharing experimentation intensity equally between them. Let $(\beta^*_{ND}, \varphi^*_{ND})$ denote an optimal Markov strategy when projects cannot be discarded individually.

Lemma 1. *Consider $(p_A(0), p_B(0))$ and the belief path $(p_A(t), p_B(t))_t$ under optimal experimentation. If $p_A(0) > p_B(0)$, then $\beta^*_{ND}(p_A(t), p_B(t)) \in \{1, S\}$ for almost all $t \in [0, \hat{t})$, where $\hat{t}$ is such that $\hat{t} = \inf\{t : p_A(t) = p_B(t)\}$. If instead $p_A(0) = p_B(0)$, then $\beta^*_{ND}(p_A(t), p_B(t)) \in \{\frac{1}{2}, S\}$ for almost all t.*[11]

Lemma 1 mimics the Gittins index representation of the optimal experimentation policy, in which a project's belief is taken to be the index.[12] In fact, the proof of Lemma 1 is essentially a simplified version of the original 'interchange argument' in Gittins and Jones (1974) and Gittins (1979) that establishes the optimality of the Gittins index for standard bandit problems.[13] Starting with an assignment in which a project with a non-maximal Gittins index is chosen before the project with the maximal index, the argument shows that interchanging the order in which both projects are pulled, keeping expected continuations following these (random) periods of experimentation fixed, increases the experimenter's payoffs.

To fully characterise an optimal experimentation strategy, Lemma 1 needs to be augmented with optimal quitting beliefs, at which all trials cease. The next result addresses this.

Proposition 1. *When projects cannot be discarded, the following admissible Markov strategy is*

[11] All proofs are in the Appendix.

[12] A special feature of exponential bandits is that the project with the highest Gittins index is also the project with the highest belief, and hence the myopically optimal allocation is also dynamically optimal. This was first shown for discrete time Bernoulli bandits by Berry and Fristedt (1985). Their result was generalised in Banks and Sundaram (1992a), who show that dynamically optimal play is myopic for a class of two-type symmetric bandits.

[13] For a concise presentation of the original proof, see Frostig and Weiss (1999).

optimal

$$\beta^*_{ND}(p_A,p_B) = \begin{cases} 1 & \text{if } p_A > p_B \text{ and } p_A > \frac{\overline{k}+\underline{k}}{G}, \\ \frac{1}{2} & \text{if } p_A = p_B > \frac{\overline{k}+\underline{k}}{G}, \\ S & \text{if } p_A \leq \frac{\overline{k}+\underline{k}}{G}. \end{cases}$$

$$\phi^*_{ND}(p_A,p_B) = \begin{cases} (1,1) & \text{if } p_A > \frac{\overline{k}+\underline{k}}{G}. \\ (0,0) & \text{if } p_A \leq \frac{\overline{k}+\underline{k}}{G}. \end{cases}$$

Continuing experimentation is optimal as long as one project's belief is above cutoff $\frac{\overline{k}+\underline{k}}{G}$. Figure 2 illustrates belief paths consistent with the optimal experimentation policy $(\beta^*_{ND},\phi^*_{ND})$. From belief (p'_A,p'_B) with $p'_A > \frac{\overline{k}+\underline{k}}{G} > p'_B$, only project A will ever be put to trial, until belief $(\frac{\overline{k}+\underline{k}}{G},p_B)$. For these beliefs, optimal payoffs have been derived in (3) (for the case of $\underline{k} = 0$), and the optimal quitting belief follows from smooth-pasting and value-matching conditions. From belief (p_A,p_B) with $p_A > p_B > \frac{\overline{k}+\underline{k}}{G}$, experimenting with project A followed by shared experimentation until belief $(\frac{\overline{k}+\underline{k}}{G},\frac{\overline{k}+\underline{k}}{G})$ is optimal. For these beliefs, Lemma 1 implies that experimentation will cease following shared experimentation. In the Appendix, I derive optimal payoffs under shared experimentation, which take a form similar to (3). The key is to note that under shared experimentation with common belief $p_A = p_B = p$, the partial differential equation (4) can be represented as a differential equation depending only on p.

4 Optimal Experimentation with Maintenance Costs

This section allows the experimenter to discard individual projects and presents the main results of the paper. When facing maintenance costs, the experimenter must balance the funding of more promising projects against the costly management of its future research options. By discarding inactive projects, the experimenter can avoid accumulating maintenance costs. However, discarding an inactive project carries an opportunity cost, since it entails the irreversible abandonment of an option value. This tension generates an incentive to bring the option value of inactive projects forward.

4.1 When to 'Go-with-the-Loser': A Culling Rule

As a first step, I provide necessary conditions for the losing project B to be put to trial under optimal experimentation. These show that the patterns of optimal experimentation when projects can be discarded are simple, in that (*i*) if a project is ever discarded it is the losing project B, and (*ii*) it is discarded as soon as the experimenter no longer intends to put it to trial, while (*iii*) whenever project A is strictly better than project B but is not currently on trial, then it must be

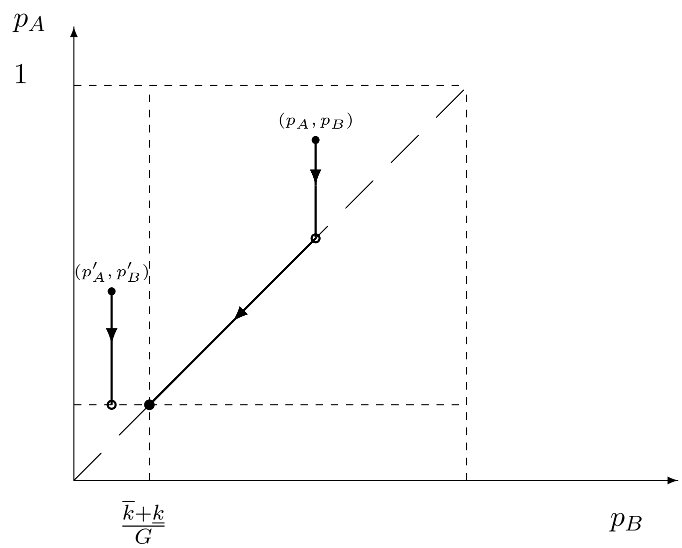

Figure 2: Optimal Experimentation Without Discarding of projects.

that project B is the only project on trial and that project A is eventually put to trial only after project B has been discarded following repeated failures. In other words, (i) and (ii) state that there is no value in 'stringing' B along without experimenting with it, only to discard is later. Furthermore, to bring the option value of the losing project B forward, the experimenter must put it to trial before it would have done so in the absence of maintenance costs, i.e., when it is still the losing project. Property (iii) establishes that experimentation with a losing project must take the form of a simple but powerful culling rule: a losing project can be given priority only in the form of a 'last chance' to produce a success. Continued failure in this period of reprieve leads to the abandonment of the project. Let (β^*, ϕ^*) denote an optimal Markov strategy.

Lemma 2. *Consider $(p_A(0), p_B(0))$ and the belief path $(p_A(t), p_B(t))_t$ under optimal experimentation.*

i. Suppose there exist $\hat{t}$ and $\epsilon > 0$ such that $\varphi^*(p_A(t), p_B(t)) \neq (1,1)$ for all $t \in [\hat{t}, \hat{t}+\epsilon)$ and $\beta^*(p_A(t), p_B(t)) \neq S$ for almost all $t \in [\hat{t}, \hat{t}+\epsilon)$. Then, without loss of generality, $\varphi^*(p_A(t), p_B(t)) = (1,0)$ for all $t \in [\hat{t}, \hat{t}+\epsilon)$.

ii. Suppose there exists $\hat{t}$ such that $\beta^*(p_A(t), p_B(t)) = 1$ for all $t \in [0, \hat{t})$ and that there exists $t' < \hat{t}$ such that $\varphi^*(p_A(t), p_B(t)) = (1,1)$ for almost all $t \in [0, t')$. Then $\varphi^*(p_A(t), p_B(t)) = (1,1)$ for all $t \in [t', \hat{t})$.

iii. Suppose that $p_A(0) > p_B(0)$ and that there exists $\hat{t} > 0$ such that $\beta^*(p_A(t), p_B(t)) \neq 1$ for *almost all* $t \in [0, \hat{t})$ *and* $\varphi^*(p_A(t), p_B(t)) = (1, 1)$ *for all* $t \in [0, \hat{t})$. *Then there exists* t^* *such that* $\hat{t} \leq t^*$, $\beta^*(p_A(t), p_B(t)) = 0$ *for almost all* $t \in [0, t^*)$, $\varphi^*(p_A(t), p_B(t)) = (1, 1)$ *for all* $t \in [0, t^*)$ *and* $\varphi^*(p_A(t^*), p_B(t^*)) = (1, 0)$.

Proving parts *i* and *ii* of Lemma 2 is simple. If the better project A were discarded before B, and B was used after having discarded A, then inverting the roles of projects A and B would increase the experimenter's payoff. If, on the other hand, project B were maintained but never used again, were project B to be discarded immediately, discoveries would occur with the same probability and maintenance costs would be avoided for a random time of positive expected length. The proof of part *iii* of Lemma 2 relies on Lemma 1, which shows that if both projects are maintained, optimal experimentation requires that the better project be used. Hence, any period of experimentation in which project B is used and project A is maintained must end by project B being discarded.

4.2 When to End a Culling Period: the Discarding Boundary

In this section, I focus on the experimenter's decision to discard project B following a culling period of experimentation. To this end, suppose that $p_A > p_B$ and that (p_A, p_B) lies in an open region of beliefs in which project B is put to trial. Then, since by Lemma 2 project B is given its 'last chance', it will be discarded in the event of failure at some belief p_B^*. As was shown in Section 2.1, the experimenter's payoff at (p_A, p_B) satisfies

$$u(p_A, p_B) = v_B(p_B; C_B(p_A)),$$

for some constant of integration $C_B(p_A)$. The experimenter's payoff at belief (p_A, p_B^*) once project B has been discarded is given by $u_A(p_A)$ and is independent of p_B. Hence value-matching and smooth-pasting conditions at the discarding belief (p_A, p_B^*) yield

$$v_B(p_B^*; C_B(p_A)) = u_A(p_A), \text{ and} \tag{8}$$

$$\frac{\partial}{\partial p_B} v_B(p_B^*; C_B(p_A)) = \frac{\partial}{\partial p_B} u_A(p_A) = 0. \tag{9}$$

Rearranging (9) yields

$$C_B(p_A) \left(\frac{1 - p_B^*}{p_B^*} \right)^{\frac{r}{G}} = \frac{G(r + \overline{k} + \underline{k})}{r(r + G)} \frac{p_B^* G}{p_B^* G + r},$$

which, along with (8), yields that p_B^* solves

$$u_A(p_A) = \frac{p_B^* G - (\overline{k} + \underline{k})}{p_B^* G + r}. \tag{10}$$

Equation (10) defines the discarding boundary. It does not determine whether project B is actually ever used when $p_A > p_B$, just when it should be discarded were it to be used. Note that the right-hand side in (10) is the payoff to a project that is known to be of type *Good* but has a success rate p_B^*G and associated experimentation cost $\overline{k} + \underline{k}$. Hence (10) states that at a cutoff belief (p_A, p_B^*) at which project B is discarded, the experimenter is indifferent between its payoff to project A in the absence of project B and a riskless project with a payoff equal to project B's flow payoff at the belief p_B^* at which it is discarded.

Note that (10) also implies that given project A with belief p_A, there is a unique candidate cutoff state (p_A, p_B^*) at which project B is discarded. Hence, define mapping $p_B^* : [0,1] \to [0,1]$ such that $p_B^*(p_A)$ is the unique solution to (10) if it exists, and is equal to p_A otherwise. Clearly, a necessary condition for project B to be put to trial before project A is that there exists belief p_A such that $p_B^*(p_A) < p_A$. This occurs whenever, for fixed p_A, there exits p_B such that $u_A(p_A) < \frac{p_B^*G-(\overline{k}+\underline{k})}{p_B^*G+r}$. To this end, consider the mapping $p_B \mapsto \frac{p_BG-(\overline{k}+\underline{k})}{p_BG+r}$. It is straightforward to verify that this mapping is increasing and concave. Hence, for fixed p_A, the inequality $u_A(p_A) \leq \frac{p_BG-(\overline{k}+\underline{k})}{p_BG+r}$ is easiest to satisfy for $p_B = p_A$. Note that

$$\begin{aligned}\lim_{p_A \nearrow 1} \left[u_A(p_A) - \frac{p_AG - (\overline{k}+\underline{k})}{p_AG + r} \right] &= \frac{G - \overline{k}}{G + r} - \frac{G - (\overline{k}+\underline{k})}{G + r} \\ &> 0. \end{aligned} \tag{11}$$

That is, as the probability that project A is of type *Good* approaches 1, the payoff to a single risky project with cost $\overline{k}$ approaches the payoff to a project known to be of type *Good* with cost $\overline{k}$ and success rate G. This dominates the payoff to a project known to be of type *Good* with cost $\overline{k} + \underline{k}$ and success rate G. In other words, the experimenter has a strict incentive to discard project B when project A is almost sure to be good. Furthermore,

$$\begin{aligned}\lim_{p_A \searrow \frac{\overline{k}+\underline{k}}{G}} \left[u_A(p_A) - \frac{p_AG - (\overline{k}+\underline{k})}{p_AG + r} \right] &= v_A(\frac{\overline{k}+\underline{k}}{G}; \tilde{C}_A) \\ &< 0. \end{aligned} \tag{12}$$

That is, as p_A approaches the quitting belief $\frac{\overline{k}+\underline{k}}{G}$ (for a risky project with cost $\overline{k} + \underline{k}$), the payoff to a project known to be of type *Good* with cost $\overline{k} + \underline{k}$ and success rate p_AG approaches 0, while the payoff to a risky project with cost $\overline{k}$ is strictly positive, since its own quitting belief is $\frac{\overline{k}}{G}$. That is, contrary to the results of Proposition 1, the experimenter will never reach the quitting belief $\frac{\overline{k}+\underline{k}}{G}$ with both projects maintained since at belief $(\frac{\overline{k}+\underline{k}}{G}, \frac{\overline{k}+\underline{k}}{G})$, the experimenter has a strict incentive to discard only a single project. (11) shows that belief $(1,1)$ must lie outside the discarding boundary, while (12) shows the same for belief $(\frac{\overline{k}+\underline{k}}{G}, \frac{\overline{k}+\underline{k}}{G})$. Thus, if project B is ever used, beliefs must lie 'between' $(\frac{\overline{k}+\underline{k}}{G}, \frac{\overline{k}+\underline{k}}{G})$ and $(1,1)$.

The simple necessary condition from above for the optimality of a culling phase, that there exits a belief p_A such that $p_B^*(p_A) < p_A$, is also sufficient to guarantee the existence of a set of beliefs with positive Lebesgue measure in which project B is put to trial before project A. The following proposition contains the main result of the paper, that not only does the optimal management of options take the form of a culling rule, but that such a rule is indeed optimal whenever it is not the case that maintenance costs are so high that the losing project is always immediately discarded.

Proposition 2. *One of the two following cases must obtain. Either*

i. $u_A(p_A) > \frac{p_A G - (\overline{k}+\underline{k})}{p_A G + r} > 0$ *for all* p_A*, and for almost all* (p_A, p_B)*,* $\varphi^*(p_A, p_B) = (1, 0)$*, or*

ii. there exist $\overline{p}_A > \underline{p}_A$ *such that* $u_A(p_A) \le \frac{p_A G - (\overline{k}+\underline{k})}{p_A G + r}$ *if and only if* $p_A \in [\underline{p}_A, \overline{p}_A]$*. Then for almost all* (p_A, p_B) *with* $\varphi^*(p_A, p_B) = (1, 1)$*,* $\beta^*(p_A, p_B) = 0$ *only if* $p_A \in [\underline{p}_A, \overline{p}_A]$ *and* $p_B \in [p_B^*(p_A), p_A]$*. Furthermore, the set* $\{(p_A, p_B) : \varphi^*(p_A, p_B) = (1, 1)$ *and* $\beta^*(p_A, p_B) = 0\}$ *has positive Lebesgue measure.*

Figure 3 illustrates the discarding boundary when the condition of part *ii* of Proposition 2 obtains. Define

$$\mathcal{P}_M = \{(p_A, p_B) : p_A \ge p_B, p_B \ge p_B^*(p_A)\},$$

which is the set of beliefs which is inside the discarding boundary. That is, $\mathcal{P}_M$ is the maintenance region, the set of beliefs inside which project B is never discarded. Further define

$$\mathcal{P}_D = \{(p_A, p_B) : p_A \ge p_B\} \setminus \mathcal{P}_M,$$

which is the set of beliefs outside the discarding boundary. This is the discarding region, in which project B can be discarded immediately or maintained but never put to trial. It is easily verified that the boundary separating $\mathcal{P}_M$ from $\mathcal{P}_D$ is concave. From state (p_A, p_B), if it is optimal to experiment with project B, then B must be put to trial until $(p_A, p_B^*(p_A))$, after which B is discarded and A must be used until $p_A^* = \frac{\overline{k}}{G}$, the quitting belief with a single risky project.

Part *ii* of Proposition 2 states that there exist beliefs for which a culling rule for project B is optimal whenever project B is not always immediately discarded. The set $\mathcal{P}_B$ in Figure 3 illustrates the beliefs for which the argument in the proof applies, which are those beliefs close to $(\underline{p}_A, \underline{p}_A)$ and $(\overline{p}_A, \overline{p}_A)$, the boundary beliefs of $\mathcal{P}_M$ on the 45-degree line. For any beliefs (p_A, p_B), the payoff to putting project A on trial (or to shared experimentation) and maintaining B is at most the payoff to using a project known to be of type *Good* with success rate $p_A G$ and experimentation cost $\overline{k} + \underline{k}$. However, near $(\overline{p}_A, \overline{p}_A)$, discarding project B yields a payoff close to the payoff to a project known to be of type *Good* with success rate $\overline{p}_A G$ but reduced experimentation cost $\overline{k}$. Hence near $(\overline{p}_A, \overline{p}_A)$, discarding project B yields strictly higher payoffs than either using project A (or shared experimentation). Yet, for beliefs strictly inside $\mathcal{P}_M$, using project B until the discarding boundary yields strictly higher payoffs than discarding it. The same

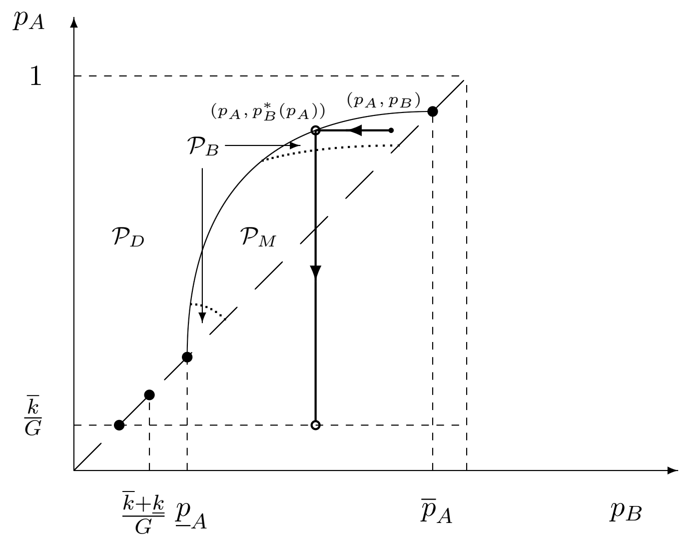

Figure 3: Discarding Boundary.

argument applies around $(\underline{p}_A, \underline{p}_A)$. Intuitively, around $(\overline{p}_A, \overline{p}_A)$ and $(\underline{p}_A, \underline{p}_A)$, the experimenter has already decided that it no longer wishes to maintain both projects in the long term. If a single project is to be maintained it should be project A. However, the option represented by project B has enough value that the experimenter wants to exploit it before discarding it, which means that project B must take precedence over project A.

4.3 Optimal Experimentation Policy

Lemmas 1 and 2 pin down the possible dynamics of trial allocations among projects. Given $p_A(0) > p_B(0)$, either *i*) project B is discarded immediately, *ii*) project B is put to trial until it is discarded in favour of project A or *iii*) project A is put to trial until either a switch to B occurs and B is put to trial until it is discarded, or belief $p_A(t)$ drops to $p_B(0)$, which leads to shared experimentation. By Proposition 2, a period of shared experimentation is always followed by a culling period for project B. In this final section, I complete the characterisation of optimal experimentation by showing when it entails these various patterns. The arguments involve backward induction from the possible endpoints of experimentation dynamics.

A useful starting point is to focus on beliefs on the 45-degree line. Lemma 2 shows that if $p_A = p_B$, then either there is shared experimentation or project B enters a culling period. The next lemma addresses the question of how many exit points from shared experimentation to project B can co-exist and shows that only a single belief $(\underline{p}, \underline{p})$ can satisfy both the value-matching and smooth-pasting conditions associated to such an exit.

Lemma 3. *Suppose there exists $p' < p''$ such that $\beta^*(p,p) = \frac{1}{2}$ for almost all $p \in [p', p'']$. Then there exists $\underline{p}$ and $\overline{p}$ such that $\underline{p} \le p'$, $p'' \le \overline{p}$ and $\beta^*(p,p) = \frac{1}{2}$ for almost all $p \in P$ if and only if $P \subset [\underline{p}, \overline{p}]$.*

For some $p > \underline{p}_A$, let $v_{AB}(p; C_{AB}(\underline{p}_A))$ be the payoff to shared experimentation at (p,p), with constant of integration $C_{AB}(p)$ capturing the effect on payoffs of moving to project B at belief $(\underline{p}_A, \underline{p}_A)$.[14] Lemma 3 implies that if the belief $\underline{p}$, which is derived explicitly in the Appendix, is such that $\underline{p} \in (\underline{p}_A, \overline{p}_A)$ and if $v_{AB}(p; C_{AB}(\underline{p})) > v_B(p; C_B(p))$ for a set of beliefs (p,p) such that $p \in (\underline{p}, \underline{p} + \epsilon]$ for some $\epsilon > 0$, then there exists $\overline{p} \in (\underline{p}, \overline{p}_A)$ such that $\beta^*(p,p) = \frac{1}{2}$ for almost all $p \in (\underline{p}, \overline{p})$ and $\beta^*(p,p) = 0$ for almost all $p \in [\underline{p}_A, \underline{p}] \cup [\overline{p}, \overline{p}_A]$. That is, optimal experimentation calls for shared experimentation only for those beliefs (p,p) with $p \in (\underline{p}, \overline{p})$. Of course, the set of beliefs for which shared experimentation is optimal can be empty.

This simple characterisation of shared experimentation allows the definition of two sets of beliefs, $\mathcal{P}_B \subset \mathcal{P}_M$ and $\mathcal{P}_A$, which will correspond the the regions of the state space in which projects A and B are used under optimal experimentation. The details, which follow from a backwards induction argument, are left to the Appendix.

Proposition 3. *When projects can be discarded, the following admissible Markov strategy is optimal.*

$$\beta^*(p_A, p_B) = \begin{cases} 0 & \text{if } (p_A, p_B) \in \mathcal{P}_B, \\ 1 & \text{if } (p_A, p_B) \in \mathcal{P}_A, \\ \frac{1}{2} & \text{if } p_A = p_B = p \text{ and } p \in (\underline{p}, \overline{p}), \\ S & \text{otherwise.} \end{cases}$$

$$\beta^*(p_A) = \begin{cases} 1 & \text{if } p_A \ge \frac{\overline{k}}{G} \\ S & \text{otherwise.} \end{cases} \qquad \beta^*(p_B) = \begin{cases} 0 & \text{if } p_B \ge \frac{\overline{k}}{G} \\ S & \text{otherwise.} \end{cases}$$

$$\varphi^*(p_A, p_B) = \begin{cases} (1,1) & \text{if } (p_A, p_B) \in \mathcal{P}_A \cup \mathcal{P}_B, \\ (1,0) & \text{otherwise.} \end{cases}$$

$$\varphi^*(p_A)) = \begin{cases} 1 & \text{if } p_A \ge \frac{\overline{k}}{G}, \\ 0 & \text{otherwise.} \end{cases} \qquad \varphi^*(p_B) = \begin{cases} 1 & \text{if } p_B \ge \frac{\overline{k}}{G}, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 4 provides an illustration of the sets $\mathcal{P}_A$ and $\mathcal{P}_B$. The figure as drawn assumes that $\mathcal{P}_A$ is convex, which need not necessarily be the case. However, note that the boundary between sets $\mathcal{P}_A$ and $\mathcal{P}_B$ must always be downward-sloping, else this would violate Lemma 2.

[14] The expression fo v_{AB} is derived in the Appendix in the proof of Proposition 1.

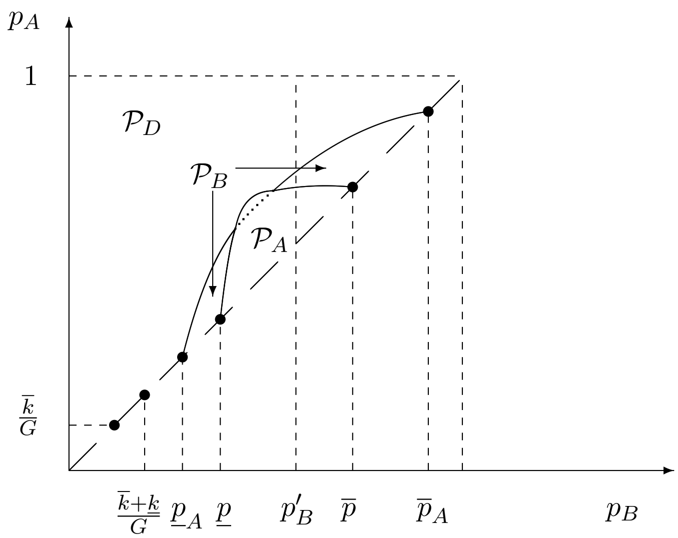

Figure 4: Optimal Experimentation Policy.

When shared experimentation occurs on the 45-degree line, which by continuity implies that $\mathcal{P}_A$ is nonempty, the reversal of the 'stay-with-the-winner' property exhibits a noteworthy non-monotonicity: for fixed p_B, 'going-with-the-loser' is optimal only if p_A is neither too high nor too low. To see this, fix p'_B such that there exist $p'_A > p''_A$ such that putting project B to trial is optimal in state (p'_A, p_B) and experimenting with project A until shared experimentation is optimal at (p''_A, p_B). If p_A is much larger than p'_A, that is, if project A is believed to be of type *Good* and hence to succeed quickly with high probability, then it is best for the experimenter to discard project B immediately and exploit project A. 'Going-with-the-loser' is optimal only for intermediate beliefs p_A that include p'_A and are no lower than p''_A. In these cases project A still has a clear advantage over project B. Optimal experimentation in the absence of maintenance costs would put project A to trial until its belief dropped to p_B, after which experimentation would be shared. However, project A is both (i) not thought likely to succeed fast enough to dwarf the option value of project B but (ii) of sufficient quality that paying to maintain the option value of project B is too costly, since this value can be realised only after project A has failed for a long time. If instead p_A is between p''_A and p_B, and hence beliefs p_A and p_B are close to each other, it is still be optimal to order project development projects as though there were no maintenance costs. In this case, both discarding project B immediately or giving it its 'last chance' is too costly, since the realisation of its option value is not so far away and project A is not the clear-cut superior project.

5 Extensions

5.1 A Culling Rule with More than Two Risky Projects

While the characterisation of the optimal experimentation policy becomes more involved, no new conceptual difficulties arise if the experimenter has more than two risky projects. The key is that experimentation dynamics following the choice of a non-winning project are qualitatively similar to those uncovered by Lemma 2 for the two-project case. Consider the case in which there are three risky projects, A, B and C, with $p_A \geq p_B \geq p_C$. Generalising the argument to the case in which there are even more risky projects is straightforward. The following result extends the culling rule to the three-project environment by showing that if it is ever optimal to experiment with the 'middle' project B, then experimentation can proceed to the winning project A only after both non-winning projects B and C have been discarded. That is, the 'last chance' extended by the experimenter applies not only to project B but to all projects ranked lower than B. That is, optimal experimentation will either put the winning project to trial or enter a targeted culling phase in which it puts to trial and then discards all sufficiently poor projects.

Proposition 4. *Consider* $(p_A(0), p_B(0), p_C(0))$ *and the belief path* $(p_A(t), p_B(t), p_C(t))_t$ *under optimal experimentation. Suppose that there exists* $\hat{t} > 0$ *such that* $\beta^*(p_A(t), p_B(t), p_C(t)) = (0, 1)$ *and* $\varphi^*(p_A(t), p_B(t), p_C(t)) = (1, 1, 1)$ *for almost all* $t \in [0, \hat{t})$*. Then there exists* $t^* > \hat{t}$ *such that (i)* $\beta^*(p_A(t), p_B(t), p_C(t)) \neq (1, 0)$ *and* $\varphi^*(p_A(t), p_B(t), p_C(t)) \in \{(1, 1, 1), (1, 1, 0)\}$ *for almost all* $t \in [\hat{t}, t^*)$*, while (ii)* $\beta^*(p_A(t), p_B(t), p_C(t)) = (1, 0)$ *and* $\varphi^*(p_A(t), p_B(t), p_C(t)) \in \{(1, 0, 0), (0, 0, 0)\}$ *for almost all* $t > t^*$*.*

Proposition 4 and with my characterisation of the two-project case could lead, through an induction argument, to a full, although tedious, characterisation of optimal experimentation with three (and then more) projects.

5.2 Complementary Projects

I have assumed that the outcomes of the projects are perfect substitutes in that the experimenter cares only about success on a single project. An alternative assumption is that success on a given project retires that project but the experimenter obtains a payoff of 1 from all projects that succeed. This section shows that the result of Proposition 2, that the set of beliefs for which it is optimal to 'go-with-the-loser' is non-negligible, continues to hold with complementary projects. Clearly, Lemmas 1 and 2 can be derived in this version of the model. It is straighforward to show that the optimal payoff in a region in which project A is put to trial and project B is maintained must follow the following version of (6)

$$u_A(p_A; p_B)[r + p_A G] + u'_A(p_A; p_B) G p_A (1 - p_A) = p_A G[1 + u_B(p_B)] - (\overline{k} + \underline{k}). \tag{13}$$

By replicating the arguments of Section 4.2, the relevant version of (10) , which determines the discarding boundary, can be shown to be

$$u_A(p_A) = \frac{p_B^* G[1 + u_A(p_A)] - (\bar{k} + \underline{k})}{G p_B^* + r}. \tag{14}$$

Since $u_A(p_A) > 0$ whenever project A is maintained, (14) show that the discarding region $\mathcal{P}_D$ is larger when projects are complimentary. Intuitively, holding on to project B longer is advantageous when sucesses can be accumulated. One difference with the discarding boundary (10) when projects are perfect substitutes is that, again intuitively, project B need not be discarded when p_A is close to 1. However, as

$$\lim_{p_A \searrow \frac{\bar{k}}{G}} \left[u_A(p_A) - \frac{p_A G[1 + u_A(p_A)] - (\bar{k} + \underline{k})}{p_A G + r} \right] < 0,$$

it must be that there exists a boundary belief $(\underline{p}_A, \underline{p}_A)$ with $\underline{p}_A > \frac{\bar{k}}{G}$ in the maintenance region $\mathcal{P}_M$. In other words, complementary projects can eliminate the incentive to 'go-with-the-loser' when the experimenter is optimistic about both projects, but 'going-with-the-loser' always benefits an experimenter that is sufficiently pessimistic about both projects. This allows the application of the argument in the proof of Proposition 2. That is, when discarding project B in favour of project A, the experimenter is guaranteed the payoff to a project known to be of type *Good* that succeeds at rate $p_B^* G[1+u_A(p_A)]$. Meanwhile, the experimenter's payoff from experimenting with project A and maintaining project B yields stricly less than the payoff to a project known to be of type *Good* that succeeds at rate $p_A G[1 + u_B(p_B)]$. The result of Proposition 2 then applies for p_B^* sufficiently close to p_A, that is, for beliefs in the region of $(\underline{p}_A, \underline{p}_A)$.

6 Conclusion

The standard approach to experimentation has been to assume that when currently occupied by other projects, keeping the option of researching various alternatives at later dates is costless. However, that keeping options open can involve maintenance costs is natural in many settings. This paper shows that such costs generate new trade-offs for experimenters by giving them incentives to manage the timing of the realisation of inactive alternatives' option values and have important implications for optimal experimentation policies.

While I have focused on the simple and tractable exponential bandit problem, it is not unreasonable to expect that my main arguments extend to more general bandit settings. Note that more generally, the arguments used in the paper are based on finite backwards induction, where the recursion is on the set of maintained projects. At each step of the recursion, the arguments rely on maintained projects' Gittins indices. This is made clear by the common structure of Lemmas 1 and 2 and the original 'interchange argument' of Gittins (1979) that establishes the optimality of index policies in standard bandit problems. Investigating the relationship between the idea of maintenance costs and general bandits is an interesting avenue for future work. This could in turn allow the model to address economic applications other than research and development, which is particularly well captured by the exponential bandit framework.

References

Banks, J. and R. Sundaram (1992a). A class of bandit problems yielding myopic optimal strategies. *Journal of Applied Probability 29*(3), 625–632.

Banks, J. and R. Sundaram (1992b). Denumerable-armed bandits. *Econometrica 60*(5), 1071–1096.

Banks, J. and R. Sundaram (1994). Switching costs and the Gittins index. *Econometrica 62*(3), 687–694.

Bergemann, D. and U. Hege (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance 22*(6-8), 703–735.

Bergemann, D. and J. Välimäki (2001). Stationary multi-choice bandit problems. *Journal of Economic Dynamics and Control 25*(10), 1585–1594.

Bergemann, D. and J. Välimäki (2006). Bandit Problems. *Cowles Foundation Discussion Papers*.

Berry, D. and B. Fristedt (1985). *Bandit problems.* London: Chapman and Hall London.

Bolton, P. and C. Harris (1999). Strategic experimentation. *Econometrica*, 349–374.

Bonatti, A. and J. Hörner (2009). Collaborating.

Bowen, M. (1987). The escalation phenomenon reconsidered: Decision dilemmas or decision errors? *Academy of Management Review 12*(1), 52–66.

Fleming, W. and R. Rishel (1975). *Deterministic and stochastic optimal control.* Berlin: Springer-Verlag.

Frostig, E. and G. Weiss (1999). Four proofs of Gittins' multiarmed bandit theorem. *Applied Probability Trust*, 1–20.

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology 75*(6), 728–731.

Gittins, J. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, 148–177.

Gittins, J. and D. Jones (1974). A dynamic allocation index for the sequential design of experiments. *Progress in statistics 241*, 266.

Hörner, J. and L. Samuelson (2009). Incentives for Experimenting Agents. *Cowles Foundation Discussion Papers*.

Jun, T. (2004). A survey on the bandit problem with switching costs. *de Economist 152*(4), 513–541.

Keller, G. and S. Rady (2009). Strategic Experimentation with Poisson Bandits. *Discussion Papers in Economics*.

Keller, G., S. Rady, and M. Cripps (2005). Strategic experimentation with exponential bandits. *Econometrica 73*(1), 39–68.

Klein, N. and S. Rady (2008). Negatively Correlated Bandits. *Discussion Papers in Economics*.

McAfee, R., H. Mialon, and S. Mialon (2010). Do Sunk Costs Matter? *Economic Inquiry 48*(2), 323–336.

Staw, B. (1981). The escalation of commitment to a course of action. *Academy of management Review 6*(4), 577–587.

Staw, B. and J. Ross (1987). Knowing when to pull the plug. *Harvard Business Review 65*(2), 68–74.

Strulovici, B. (2009). Learning While Voting: Determinants of Collective Experimentation. *Working paper*.

A Appendix

Proof of Lemma 1. Suppose that $p_A(0) > p_B(0)$, and consider the belief path under optimal experimentation $(p_A(t), p_B(t))_t$. The first step is to show that if there exists $\hat{t} > 0$ and $\hat{T} \subset [0, \hat{t})$ such that $\hat{T}$ has positive Lebesgue measure and $\beta^*(p_A(t), p_B(t)) \neq 1$ for all $t \in \hat{T}$, then $\beta^*(p_A(t), p_B(t)) = 0$ for almost all $t \in \hat{T}$. Suppose instead that $\beta^*(p_A(t), p_B(t)) = \alpha(t) \in (0, 1)$ for all $t \in \hat{T}$. Let $T_A = \int_0^{\hat{t}} \alpha(t)dt$. By assumption, $T_A \in (0, \hat{t})$.

Given allocation $\alpha(t)$ for $t \in [0, \hat{t})$, and initial beliefs $(p_A(0), p_B(0))$, solving the differential equation for the evolution of beliefs yields that

$$p_A(t) = \frac{1}{1 + \frac{1-p_A(0)}{p_A(0)} e^{H \int_0^t \alpha(s)ds}}, \text{ and}$$

$$p_B(t) = \frac{1}{1 + \frac{1-p_B(0)}{p_B(0)} e^{H \int_0^t (1-\alpha(s))ds}}.$$

Belief $p_A(t)$ depends only on the cumulative experimentation on project A up to time t, $\int_0^t \alpha(s)ds$, and not on when this experimentation occurred within the interval $[0,t]$.

Consider an alternative admissible Markov assignment $\hat{\beta}$ such that

$$\hat{\beta}(p_A(t), p_B(t)) = \begin{cases} 1 & \text{for all } t \in (0, \hat{t} - T_A], \\ 0 & \text{for all } t \in (\hat{t} - T_A, \hat{t}), \end{cases}$$

with $\hat{\beta} = \beta^*$ otherwise. Then $(\hat{p}_A(\hat{t}), \hat{p}_B(\hat{t})) = (p_A(\hat{t}), p_B(\hat{t}))$, where $(\hat{p}_A(t), \hat{p}_B(t))_t$ is the belief path associated with $\hat{\beta}$. Hence, the payoffs following $\hat{t}$ are the same under both assignments. That is, $v(\hat{\beta}, \varphi^*; p_A(\hat{t}), p_B(\hat{t})) = u(p_A(\hat{t}), p_B(\hat{t}))$. Furthermore, conditional on $(p_A(0), p_B(0))$, the probability that no success occurs until $\hat{t}$ is the same under β and $\hat{\beta}$.

Let τ_{β^*} (respectively $\tau_{\hat{\beta}}$) be the random arrival time of a success under assignment β^* (respectively $\hat{\beta}$) in time interval $[0, \hat{t}]$. Then $Pr[\tau_{\hat{\beta}} \leq t | p_A(0), p_B(0)] > Pr[\tau_{\beta^*} \leq t | p_A(0), p_B(0)]$ for all $t \in (0, \hat{t})$, that is, $\tau_{\hat{\beta}}$ is higher than τ_{β^*} in the sense of first order stochastic dominance. By discounting, the experimenter's payoff is decreasing in the arrival time of a success, and hence $\hat{\beta}$ yields a strictly higher expected payoff than β^* in $[0, \hat{t}]$, or

$$\int_0^{\hat{t}} [1 - (\overline{k} + \underline{k})] e^{-r\tau_{\hat{\beta}}} \hat{\mu}(d\tau_{\hat{\beta}}) > \int_0^{\hat{t}} [1 - (\overline{k} + \underline{k})] e^{-r\tau_{\beta^*}} \mu^*(d\tau_{\beta^*}),$$

where $\hat{\mu}$ and μ^* are the distributions of $\tau_{\hat{\beta}}$ and τ_{β^*}, respectively. Hence,

$$\begin{aligned} v(\hat{\beta}; p_A(\hat{t}), p_B(\hat{t})) &= \int_0^{\hat{t}} [1 - (\overline{k} + \underline{k})] e^{-r\tau_{\hat{\beta}}} \hat{\mu}(d\tau_{\hat{\beta}}) + Pr[\tau_{\hat{\beta}} > \hat{t} | p_A(0), p_B(0)) u(p_A(\hat{t}), p_B(\hat{t})] \\ &> \int_0^{\hat{t}} [1 - (\overline{k} + \underline{k})] e^{-r\tau_{\beta^*}} \mu^*(d\tau_{\beta^*}) + Pr[\tau_{\beta^*} > \hat{t} | p_A(0), p_B(0)) u(p_A(\hat{t}), p_B(\hat{t})] \\ &= u(p_A(0), p_B(0)), \end{aligned}$$

a contradiction. Hence it must be that $\alpha(t) = 0$ for almost all $t \in [0, \hat{t}]$.

That is, the previous argument shows that if project A is not used, it must be that project B is used exclusively. Since in that case $p_A(t) > p_B(t)$ for all $t > 0$, the previous argument also ensures that project B is used until experimentation ceases, which must occur at time t^* such that $p_B(t^*) = \frac{\overline{k} + \underline{k}}{G}$. By mimicking this strategy with project A instead of B, that is, using project A until belief $p_A^* = \frac{\overline{k} + \underline{k}}{G}$ and then moving permanently to S, the experimenter's payoff at time 0 would be higher. That is, consider alternative strategy $\hat{\beta}$ such that

$$\hat{\beta}(p_A, p_B) = \begin{cases} 1 & \text{if } p_A > \frac{\overline{k} + \underline{k}}{G} \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} v(\beta; p_A(0), p_B(0)) &= v_A(p_A(0); C_A) \\ &\geq v_B(p_B(0, C_A)) \\ &= u(p_A(0), p_B(0)), \end{aligned}$$

a contradiction. C_A is the constant of integration for the optimal stopping problem with a single risky project and direct cost $\overline{k} + \underline{k}$. Hence, it must be that $\alpha(t) = 1$ for all t such that $p_A(t) > p_B(t)$.

The same argument can be applied if $p_A(0) = p_B(0)$ to show that experimentation is shared until it ceases, i.e., $\beta^*(p_A(t), p_B(t)) = \frac{1}{2}$ for all t such that $\varphi(p_A(t), p_B(t))$.

□

Proof of Proposition 1. To obtain an expression for optimal payoffs under shared experimentation, note that under the assumption that $p_A = p_B = p$ and that $\beta(p,p) = \frac{1}{2}$ for all beliefs p greater than some quiting belief p^*, the optimal payoff u must satisfy $\frac{\partial}{\partial p_A} u = \frac{\partial}{\partial p_B} u$. Define $u_{AB}(p) \equiv u(p,p)$, then it follows that $u'_{AB}(p) = 2\frac{\partial}{\partial p_A} u(p,p)$ and u_{AB} solves

$$u_{AB}(r + pG) = pG - (\overline{k} + \underline{k}) - \frac{1}{2} Gp(1-p)u'_{AB}.$$

The differential equation (15) has solution

$$u_{AB}(p) = \tilde{C}_{AB} \left(\frac{1-p}{p}\right)^{\frac{2r}{G}} (1-p)^2 + p^2 \frac{G - (\overline{k} + \underline{k})}{r + G} + 2p(1-p) \frac{\frac{G}{2} - (\overline{k} + \underline{k})}{r + \frac{G}{2}} - (1-p)^2 \frac{\overline{k} + \underline{k}}{r}. \quad (15)$$

If optimal experimentation leads from shared experimentation to the safe project at belief p^*, the constant of integration $\tilde{C}_{AB}$ and cutoff belief $p^* = \frac{\overline{k}+\underline{k}}{G}$ are determined by value-matching and smooth-pasting conditions.

$$\begin{aligned} &u_{AB}(p^*) = 0, \text{ and} \\ &u'_{AB}(p^*) = 0. \end{aligned}$$

What remains to be shown in the text is that the Markov strategy $(\beta^*_{ND}, \phi^*_{ND})$ is admissible and that the value function $w(p_A, p_B)$ inferred from strategy $(\beta^*_{ND}, \phi^*_{ND})$ solves the Bellman equation

$$\begin{aligned} ru(p_A, p_B) = \max \Bigg\{ 0,\ & p_A G - (\overline{k} + \underline{k}) - \frac{\partial u(p_A, p_B)}{\partial p_A} G p_A (1 - p_A) - u(p_A, p_B) G p_A, \\ & p_B G - (\overline{k} + \underline{k}) - \frac{\partial u(p_A, p_B)}{\partial p_B} G p_B (1 - p_B) - u(p_A, p_B) G p_B \Bigg\}. \end{aligned} \quad (16)$$

Clearly, given any t^*, t^{**} and t' such that $0 \le t^* \le t^{**}$ and $t' \ge 0$, any strategy (α, ϕ) of the form

$$\alpha(t) = \begin{cases} 1 & \text{if } t < t^*, \\ \frac{1}{2} & \text{if } t \in [t^*, t^{**}), \\ S & \text{if } t \ge t^{**}, \end{cases}$$
$$\phi(t) = \begin{cases} (1,1) & \text{if } t < t', \\ (0,0) & \text{if } t \ge t', \end{cases} \tag{17}$$

is admissible. Furthermore, given any (p_A, p_B), there exist t^*, t^{**} and t' such that $0 \le t^* \le t^{**}$ and $t' \ge 0$ such that a strategy (α, ϕ) defined as in (17) is such that

$$\begin{aligned} \alpha(t) &= \beta^*_{ND}(p_A(t), p_B(t), \phi_A(t), \phi_B(t)), \\ \phi_A(t) &= \varphi^*_{A,ND}(p_A(t), p_B(t), \phi_A(t), \phi_B(t)), \\ \phi_B(t) &= \varphi^*_{B,ND}(p_A(t), p_B(t), \phi_A(t), \phi_B(t)), \end{aligned}$$

and hence Markov strategy $(\beta^*_{ND}, \varphi^*_{ND})$ is admissible.

Let $w(p_A, p_B)$ be the value function inferred from strategy $(\beta^*_{ND}, \phi^*_{ND})$. Consider state (p_A, p_B) such that $p_A > p_B > \frac{\overline{k}+\underline{k}}{G}$. Then

$$w(p_A, p_B) = C_A(p_B) \left(\frac{1-p_A}{p_A} \right)^{\frac{r}{G}} (1-p_A) + p_A \frac{G - (\overline{k}+\underline{k})}{r+G} - (1-p_A) \frac{\overline{k}+\underline{k}}{r},$$

with the constant of integration determined at the switch to shared experimentation when $p_A = p_B$. It is easy to see that $w(p_A, p_B) > w(p_B, p_B) > 0$, and hence it is optimal to continue experimentation for all $p_A > p_B$. The derivative of the third term of (16) with respect to p_A is $-\frac{\partial}{\partial p_A} w(p_A, p_B) G p_B < 0$. Since $w(p_A, p_B)$ is increasing in p_A, so is the second term of (16). If $p_A \to p_B$, the difference in the two terms is given by $G p_B (1-p_B) \left[\frac{\partial w(p_A,p_B)}{\partial p_A} - \frac{\partial w(p_A,p_B)}{\partial p_B} \right] \Big|_{p_A = p_B} = 0$ (by smooth-pasting since at (p_B, p_B) a transition occurs to shared experimentation). Hence, the second term of (16) is larger than the third term for all $p_A > p_B$, as required. For (p_A, p_B) such that $p_A > \frac{\overline{k}+\underline{k}}{G} > p_B$, then $w(p_A, p_B) = u_A(p_A)$, while for (p_A, p_B) such that $p_A = p_B > \frac{\overline{k}+\underline{k}}{G}$, $w(p_A, p_B) = u_{AB}(p_A)$. In both these cases, it is easy to see that value-matching and smooth-pasting imply that $w(p_A, p_B) > 0$ whenever $p_A > \frac{\overline{k}+\underline{k}}{G}$, as desired.

□

Proof of Lemma 2. For part i, suppose there exists $\hat{t}$ and $\epsilon > 0$ such that $\varphi^*(p_A(t), p_B(t)) \neq (1,1)$ and $\beta^*(p_A(t), p_B(t)) \neq S$ for almost all $t \in [\hat{t}, \hat{t}+\epsilon)$. Then one project is discarded on the equilibrium path. Let $t^* = \inf\{t < \hat{t} : \varphi^*(p_A(t), p_B(t)) \neq (1,1)\}$. If $\varphi^*(p_A(t^*), p_B(t^*)) = (0,1)$, then since $\beta^*(p_A(t^*), p_B(t^*)) \neq S$ for almost all $t \in [\hat{t}, \hat{t}+\epsilon)$, it must be that $\beta^*(p_A(t^*), p_B(t^*)) = 0$

for almost all $t \in [\hat{t}, \hat{t}+\epsilon)$. Consider a Markov strategy (β', φ') such that

$$\varphi'(p_A, p_B) = (1,0) \qquad \text{for all } (p_A, p_B) \text{ such that } \varphi^*(p_A, p_B) = (0,1),$$
$$\beta'(p_A(t), p_B(t)) = 1 \qquad \text{for all } t > t^* \text{ for which } \beta^*(p_A(t), p_B(t)) = 0,$$

with $(\beta', \varphi') = (\beta^*, \varphi^*)$ otherwise. Under (β', φ'), $p'_A(t) \geq p_B(t)$ for all $t > t^*$ by the assumption of symmetric strategies, and hence for all $t > t^*$ such that $\beta^*(p_A(t), p_B(t)) = 0$,

$$\begin{aligned} v(\beta', \varphi'; p_A(t), p_B(t)) &= v_A(p_A(t); \tilde{C}_A) \\ &\geq v_B(p_B(t); \tilde{C}_A) \\ &= v_B(p_B(t); \tilde{C}_B) \\ &= u(p_A(t), p_B(t)). \end{aligned}$$

If the inequality is strict, this yields the required contradiction, while if it holds with equality, it is without loss of generality to discard project B instead of project A.

For part ii, suppose that there exists $\hat{t}$ such that $\beta^*(p_A(t), p_B(t)) = 1$ for almost all $t \in [0, \hat{t})$ and that there exists $t' < \hat{t}$ such that $\varphi^*(p_A(t), p_B(t)) = (1,1)$ for almost all $t \in [0, t')$, but that $\varphi^*(p_A(t''), p_B(t'')) \neq (1,1)$ for some $t'' \in (t', \hat{t})$. By part i, $\varphi^*(p_A(t''), p_B(t'')) = (1,0)$. Consider Markov strategy (β', φ') such that $\varphi'(p_A(0), p_B(0)) = (1,0)$, with $(\beta', \varphi') = (\beta^*, \varphi^*)$ otherwise. Then we can write

$$v(\beta', \varphi'; p_A(0), p_B(0)) = v_A(p_A(0); C'_A) + p_A(0)\frac{k}{r+G} + (1-p_A(0))\frac{k}{r},$$

and

$$u(p_A(t), p_B(t)) = v_A(p_A(0); C_A),$$

where the constants of integration C_A and C'_A are determined at beliefs $(p_A(t''), p_B(t''))$ at which

$$v_A(p_A(t''); C_A) = v_A(p_A(t''); C'_A) = u(p_A(t''), p_B(t'')).$$

Hence

$$C'_A = C_A - \frac{p_A(t'')\frac{k}{r+G} + (1-p_A(t''))\frac{k}{r}}{\left(\frac{1-p_A(t'')}{p_A(t'')}\right)^{\frac{r}{G}}(1-p_A(t''))},$$

and

$$\begin{aligned} v(\beta', \varphi'; p_A(0), p_B(0)) &= v_A(p_A(0); C_A) \\ &\quad - \frac{\left(\frac{1-p_A(0)}{p_A(0)}\right)^{\frac{r}{G}}(1-p_A(0))}{\left(\frac{1-p_A(t'')}{p_A(t'')}\right)^{\frac{r}{G}}(1-p_A(t''))}\left[p_A(t'')\frac{k}{r+G} + (1-p_A(t''))\frac{k}{r}\right] \\ &\quad + p_A(0)\frac{k}{r+G} + (1-p_A(0))\frac{k}{r} \\ &> v_A(p_A(0); C_A). \end{aligned}$$

The inequality follows since $p_A(t'') < p_A(0)$.

$\square$

Proof of Lemma 2. Lemma 1 and part *i* of Lemma 2 imply that if there exists $\hat{t} > 0$ such that $\beta^*(p_A(t), p_B(t)) \neq 1$ for almost all $t \in [0, \hat{t})$ and $\varphi^*(p_A(t), p_B(t)) = (1,1)$ for all $t \in [0, \hat{t})$, then it must be that $\beta^*(p_A(t), p_B(t)) = 0$ for almost all $t \in [0, \hat{t})$ and that if there exists $t^* \geq \hat{t}$ such that $\beta^*(p_A(t^*), p_B(t^*)) > 0$, then by part *i* of Lemma 2 it must be that $\varphi^*(p_A(t^*), p_B(t^*)) \in \{(1,0),(0,0)\}$. That is, if project A is not pulled it must be that project B is, and the experimenter cannot go back to project A without discarding project B. Since the experimenter must eventually discard B if p_B gets close to 0, it only remains to be shown that $\varphi^*(p_A(t^*), p_B(t^*)) = (1,0)$, that is, that the experimenter will discard B in favour of A at t^*. This follows by part *ii* of Lemma 2.

$\square$

Proof of Proposition 2. First, $u_A(p_A)$ is increasing and convex in p_A. Also, since the mapping $p_A \mapsto \frac{p_A G - (\overline{k} + \underline{k})}{p_A G + r}$ is increasing and concave, then by (11) and (12) either $u_A(p_A) > \frac{p_A G - (\overline{k} + \underline{k})}{p_A G + r} > 0$ for all p_A or there exist $\overline{p}_A > \underline{p}_A$ such that $u_A(p_A) \leq \frac{p_A G - (\overline{k} + \underline{k})}{p_A G + r}$ if and only if $p_A \in [\underline{p}_A, \overline{p}_A]$, where $\underline{p}_A$ and $\overline{p}_A$ are the only two solutions to $u_A(p_A) = \frac{p_A G - (\overline{k} + \underline{k})}{p_A G + r}$.

Now suppose that the conditions of part *ii* obtain. A first claim is that at $(\overline{p}_A, \overline{p}_A)$, discarding project B is strictly preferred to shared experimentation. By Lemma 2, if B is not discarded then $\beta^*(\overline{p}_A, \overline{p}_A) = \frac{1}{2}$ and the beliefs go down the 45-degree line until some belief (p^*, p^*), and hence the experimenter's payoffs satisfy $u(p_A, p_B) = v_{AB}(\overline{p}_A; C_{AB}(p^*))$. v_{AB} itself satisfies

$$\begin{aligned} v_{AB}(\overline{p}_A; C_{AB}(p^*)) &= \frac{\overline{p}_A G - (\overline{k} + \underline{k})}{(r + \overline{p}_A G)} - \frac{G \overline{p}_A (1 - \overline{p}_A)}{2(r + \overline{p}_A G)} v'_{AB}(\overline{p}_A; C_{AB}(p^*)) \\ &< \frac{\overline{p}_A G - (\overline{k} + \underline{k})}{(r + \overline{p}_A G)} \\ &= v_A(\overline{p}_A; \tilde{C}_A). \end{aligned}$$

Hence, by continuity, for states (p, p) with $p < \overline{p}_A$ sufficiently close to $\overline{p}_A$, discarding B is strictly preferred to shared experimentation. A very similar argument shows that discarding B is strictly preferred to using project A for an open set of states of positive Lebesgue measure (p_A, p_B) with $p_A > p_B$ sufficiently close to $(\overline{p}_A, \overline{p}_A)$. However, within the discarding boundary using project B (until the boundary) is preferred to discarding it and hence there exists an open region of positive Lebesgue measure around $(\overline{p}_A, \overline{p}_A)$ in which using project B is optimal. A very similar argument demonstrates the same result for a region around $(\underline{p}_A, \underline{p}_A)$.

$\square$

Proof of Lemma 3. By Lemma 2, once the experimenter quits shared experimentation, project B is used, then discarded and replaced with project A. Also, by Proposition 2, there exists a belief $\hat{p} > \underline{p}_A$ such that $\beta^*(p, p) = 0$ for almost all $p \in [\underline{p}_A, \hat{p}]$. Suppose there exists $p' > p''$

such that $\beta^*(p,p) = \frac{1}{2}$ for almost all $p \in [p', p'']$, and that the experimenter switches from shared experimentation to project B at belief $p^* < p''$. Hence the smooth-pasting condition at belief p^* is

$$\begin{aligned}\frac{\partial}{\partial p_B} u(p^*, p^*) &= \frac{1}{2} \cdot \frac{\partial}{\partial p} v_{AB}(p^*; C_{AB}(p^*)) \\ &= \frac{\partial}{\partial p_B} v_B(p^*, C_B(p^*)),\end{aligned}$$

which, with manipulations, yields that

$$\begin{aligned}C_{AB} = &\frac{C_B(p)}{\left(\frac{1-p}{p}\right)^{\frac{r}{G}} (1-p)} \\ &+ \frac{p\left[\frac{G-(\overline{k}+\underline{k})}{r+G} - \frac{\frac{G}{2}-(\overline{k}+\underline{k})}{r+\frac{G}{2}}\right] + (1-p)\left[\frac{\frac{G}{2}-(\overline{k}+\underline{k})}{r+\frac{G}{2}} - \frac{\overline{k}+\underline{k}}{r}\right] - \frac{G(r+\overline{k}+\underline{k})}{r(r+G)}}{\left(\frac{1-p}{p}\right)^{\frac{2r}{G}+1} \frac{Gp+r}{G}}.\end{aligned}$$

Meanwhile, the value matching condition is

$$v_{AB}(p^*; C_{AB}(p^*)) = v_B(p^*; C_B(p^*)),$$

which, with manipulations, yields that

$$C_{AB} = \frac{C_B(p)}{\left(\frac{1-p}{p}\right)^{\frac{r}{G}} (1-p)} - \frac{\frac{G^2}{2}(r + \overline{k} + \underline{k})}{\left(\frac{1-p}{p}\right)^{\frac{2r}{G}+1} r(r+G)(r+\frac{G}{2})}.$$

Together, these yield that

$$p^* = \frac{2(\overline{k} + \underline{k})(r+G)(r+\frac{G}{2})}{2(\overline{k} + \underline{k})(r+G)(r+\frac{G}{2}) + \frac{G^2}{2}(\overline{k} + \underline{k} + r)} \tag{18}$$

Clearly, $p^* \in [0,1]$ is unique. Define $\underline{p}$ to be the unique solution to (18).

□

Proof of Proposition 3. First, I construct the sets $\mathcal{P}_B$ and $\mathcal{P}_A$. In the following, assume that the conditions of Lemma 3 are met and that there exists a (unique) portion of the 45-degree line $(\underline{p}, \overline{p})$ for which shared experimentation is optimal. The arguments that follow apply in a straightforward way when this is not the case.

First, let

$$\mathcal{P}_B^1 = \left\{(p_A, p_B) \in \mathcal{P}_M : p_A \in [\underline{p}_A, \underline{p}] \cup [\overline{p}, \overline{p}_A]\right\}.$$

By Lemma 2, it must be that given $p \in [\underline{p}_A, \underline{p}] \cup [\overline{p}, \overline{p}_A]$, $\beta^*(p, p_B) = 0$ for almost all $p_B \in (p, p_B^*(p))$. That is, from the 45-degree line, if there is no shared experimentation then project B enters a culling period.

Second, consider

$$\mathcal{P}_M \setminus \mathcal{P}_B^1 = \Big\{ (p_A, p_B) \in \mathcal{P}_M : p_A \in [\underline{p}, \overline{p}], p_B \in [p_B^*(\underline{p}), \overline{p}] \Big\},$$

the set of beliefs in the maintenance region that have not been attributed to $\mathcal{P}_B^1$. By Lemma 2, from such beliefs, an optimal policy will either put project B to trial immediately until it is discarded, or put project A to trial either until beliefs reach the 45-degree or until a switch to project B occurs. Define $v_A(p_A; C_A(p_B; p_A'))$ to be the payoff to the experimenter in state $(p_A, p_B) \in \mathcal{P}_M \setminus \mathcal{P}_B^1$ were it to put project A to trial until belief $p_A' \in [\max\{p_B, \underline{p}\}, p_A)$, and then switch to project B until discarding belief $p_B^*(p_A')$. Hence the constant of integration $C_A(p_B; p_A')$ depends on the belief p_B and on the switching belief p_A', but not on p_A. Similarly, if $p_B > \underline{p}$, define $v_A(p_A; C_A^{45}(p_B))$ to be the payoff to the experimenter in state (p_A, p_B) were it to put project A to trial until it reaches the 45-degree line, after which it shares experimentation until joint belief $\underline{p}$. If $p_B \leq \underline{p}$, then define $v_A(p_A; C_A^{45}(p_B)) = v_A(p_A; C_A(p_B; \underline{p}))$. Note that $v_A(p_A; C_A(p_B; p_A')) \geq v_A(p_A; C_A(p_B; p_A''))$ if and only if $C_A(p_B; p_A') \geq C_A(p_B; p_A'')$ and $v_A(p_A; C_A(p_B; p_A')) \geq v_A(p_A; C_A^{45}(p_B))$ if and only if $C_A(p_B; p_A') \geq C_A^{45}(p_B)$. Hence if $C_A(p_B; p_A) = \max_{\{p_A' \in [\max\{p_B, \underline{p}\}, p_A]\}} C_A(p_B; p_A')$, then the experimenter has no incentive to put project A to trial. If, on the other hand, there exists a p_A' such that $C_A(p_B; p_A') > C_A(p_B; p_A)$, then the experimenter gains by staying with project A until belief p_A'. Let

$$\mathcal{P}_B^2 = \left\{ (p_A, p_B) \in \mathcal{P}_M \setminus \mathcal{P}_B^1 : \max\left\{ \max_{p' \in [\max\{p_B, \underline{p}\}, p_A]} C_A(p_B; p'), C_A^{45}(p_B)) \right\} \leq C_A(p_B; p_A) \right\},$$

and let $\mathcal{P}_B = \mathcal{P}_B^1 \cup \mathcal{P}_B^2$. Finally, let $\mathcal{P}_A^1 = \mathcal{P}_M \setminus \mathcal{P}_B$. Hence, all the beliefs in $\mathcal{P}_M$ have been attributed either to $\mathcal{P}_B$ or to $\mathcal{P}_A^1$.

Third, consider the beliefs in $\mathcal{P}_D$, those outside the discarding boundary. By Lemma 2, it must be that $\varphi^*(p_A, p_B) = (1, 0)$ for all $(p_A, p_B) \in \mathcal{P}_D$ that are not in the set $\{(p_A, p_B) : \beta^*(p_A, p_B(p_A)) = 1\}$. That is, if project B is maintained, it must be that it will not be discarded once beliefs reach the discarding boundary. Let

$$\mathcal{Q} = \Big\{ (p_A, p_B) \in \mathcal{P}_D : (p_A, p_B^*(p_A)) \in \mathcal{P}_A^1 \Big\}.$$

That is, $\mathcal{Q}$ is the set of beliefs in the discarding region such that were A to be used and B maintained until the discarding bound, B would also be maintained when beliefs cross into $\mathcal{P}_M$. For any $(p_A, p_B) \in \mathcal{Q}$, define $p_A^{**}(p_B) = \sup\{p_A : (p_A, p_B) \in \mathcal{P}_A^1\}$. Furthermore, define

$$\mathcal{P}_A^2 = \Big\{ (p_A, p_B) \in \mathcal{Q} : v_A(p_A; C_A(p_B; p_A^{**}(p_B)) > v_A(p_A; \tilde{C}_A) \Big\}.$$

Finally, let $\mathcal{P}_A = \mathcal{P}_A^1 \cup \mathcal{P}_A^2$.

What remains is to show that the strategy (β^*, φ^*) is admissible and that the value function $w(p_A, p_B)$ inferred from strategy (β^*, ϕ^*) solves the Bellman equation for the experimenter's problem. Both of these follow from an argument very similar to that for Proposition 1. In this case, however, verification of the optimality of (β^*, φ^*) is more tedious, and is omitted. □

Proof of Proposition 4. Consider $(p_A(0), p_B(0), p_C(0))$ and the belief path $(p_A(t), p_B(t), p_C(t))_t$ under optimal experimentation. Suppose that there exists $\hat{t} > 0$ such that $\beta^*(p_A(t), p_B(t), p_C(t)) = (0, 1)$ and $\varphi^*(p_A(t), p_B(t), p_C(t)) = (1, 1, 1)$ for almost all $t \in [0, \hat{t})$. By Lemma 2, if project C is ever put to trial, then it is on trial continuously until it is discarded. Hence the statement of Proposition 4 fails if there exist $t_1' < t_2' \leq t_1'' < t_2'' \leq t_1''' < t_2'''$ such that $\beta^*(p_A(t), p_B(t), p_C(t)) \notin \{(1, 0), (0, 0)\}$ for almost all $t \in (t_1', t_2') \cup (t_1''', t_2''')$, $\beta^*(p_A(t), p_B(t), p_C(t)) = (1, 0)$ for almost all $t \in (t_1'', t_2'')$ and $\varphi^*(p_A(t), p_B(t), p_C(t)) \in \{(1, 1, 1), (1, 1, 0)\}$ for all $t \in (t_1', t_2') \cup (t_1'', t_2'') \cup (t_1''', t_2''')$. Then an argument along the lines of that of Lemma 1 yields that reordering the assignment by experimenting with project A exclusively earlier increases the experimenter's payoff, yielding the required contradiction. □