Keeping Your Options Open*

Jean Guillaume Forand†

February 8, 2015

Abstract

In standard models of experimentation, the costs of project development consist of (*a*) the direct cost of running trials as well as (*b*) 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 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, optimal experimentation policies have 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 optimal experimentation policies, 'going with the loser' can be optimal: projects that are less likely to succeed are sometimes developed first.

JEL Classification: D83, D81, C61

Keywords: Experimentation, Maintenance Costs, Multi-Armed Bandits

1 Introduction

When experimentation is costly, decision-makers must choose which alternatives to actively investigate and which to leave 'on the back burner'. A firm engaged in research and development can obtain comparable product innovations through various technologies. Investing in multiple technologies simultaneously is costly, so how should the firm prioritise its allocation of funds to

*I would like to thank Li Hao, Martin Osborne and Colin Stewart for their supervision, comments and suggestions. I would also like to thank Phil Curry, Ettore Damiano, John Duggan, Francisco Gonzalez, Carolyn Pitchik, Wing Suen, Thomas Wiseman, seminar participants at Guelph, Rochester and Waterloo and participants at the CETC 2011, Fall 2010 METC, NASM 2011 SEDesign 2011 and SEDynamics 2011. Finally, the Co-Editor, Herbert Dawid, and three anonymous referees have provided excellent feedback. One referee in particular has had an exceptional and invaluable involvement with this paper, for which I am very grateful.

†Department of Economics, University of Waterloo, Hagey Hall of Humanities, Waterloo, Ontario, Canada N2L 3G1. Email: jgforand@uwaterloo.ca. Website: http://arts.uwaterloo.ca/∼jgforand

competing alternatives? Academic researchers typically have many ongoing projects of varying quality, with limited time to devote to research duties. Which projects should receive more of the researcher's attention? In standard models of experimentation, putting one alternative to trial over another entails only an implicit opportunity cost: the foregone chance of learning about the inactive alternative.[1] However, retaining the option to investigate a currently shelved alternative often involves explicit maintenance costs. Firms engaged in research and development routinely devote resources to keep open the option of developing low-priority technologies, which involves the costly upkeep of specialised equipment and paying the salaries of skilled workers or scientists that can be lost to other firms. For these firms, holding on to the rights to proprietary technology with outside market value also imposes a cost in foregone revenue. Academic researchers can bear psychological costs from lingering projects, and efforts to limit these costs may impact their allocation of time to ongoing projects.

In this paper, I present a multi-armed bandit model of experimentation in which arms (projects) that are not allocated experimentation effort impose explicit maintenance costs unless the experimenter chooses to irreversibly discard them. Specifically, I adopt the continuous-time exponential bandit framework of Keller et al. (2005) and consider a model in which two independent risky projects can be either good or bad, with only good projects eventually succeeding if developed. The experimenter's belief about a project, its updated probability that the project is good, decays as the project is put to trial and no success is observed. The experimenter selects discarding times for the projects and allocates trials among maintained projects to maximise its discounted expected payoffs: the benefits from obtaining a success on either project less total experimentation costs, which include the direct trial costs of active projects and the maintenance costs of inactive projects. Successes are perfect substitutes and a success on either project ends experimentation. As opposed to standard problems, at any time at which some project is maintained but inactive, the experimenter faces a choice between paying to keep that option open or discarding it (irreversibly) altogether. Discarding an inactive project liquidates its option value, which is realised in the event that the currently active project is deemed unpromising. To avoid destroying this option value or paying to maintain it, the experimenter has an incentive to bring it forward by altering the order of project development.

I show that optimal experimentation policies with maintenance costs entail significant departures from standard results. In the corresponding environment without maintenance costs, optimal experimentation policies are characterised by the well-known Gittins index policy.[2] In the exponential framework, this policy has the 'stay-with-the-winner' property: at any time, the project with the highest updated probability of being good (the winner) is put to trial. In the presence of maintenance costs, I show that 'going with the loser' is a robust feature of optimal experimentation: the project with the lowest updated probability of being good (the loser) can

[1] Bergemann and Välimäki (2008) survey the multi-armed bandits literature with applications to economics.

[2] To each arm is assigned an index number that depends only on the ex ante characteristics and accumulated observations of that project. The optimal experimentation selects a project among those with maximal indices.

be put to trial first. Specifically, either (*a*) maintenance costs are so high that the losing project is always discarded immediately for all initial beliefs about its quality, or (*b*) under the optimal experimentation policy the loser will be put to trial before the winning project for a non-negligible set of initial beliefs.

Under the stay-with-the-winner rule, the experimenter can avoid accumulating maintenance costs on the inactive losing project only by discarding this project. Since abandoning a project is irreversible, this involves an opportunity cost attached to the eventuality in which the losing project becomes valuable following repeated failures by the current winner. This tension generates an incentive to reverse the order of experimentation. When the experimenter prioritises the losing project, she does so through a simple culling rule. This project is granted a 'last chance' to succeed in a period of experimentation in which it is put to trial exclusively while the winner is maintained, after which it is permanently discarded. In other words, if the losing project is ever put to trial, the experimenter returns to the winning project only after the losing project has been discarded.

An outside observer that fails to take maintenance costs for inactive projects into account would conclude that the experimenter is sometimes prioritising the wrong project. Furthermore, during a culling phase the experimenter appears to not 'know when to pull the plug' and to cling to projects that have repeatedly failed to achieve results.[3] However, when inactive projects' option values are attached to a stream of maintenance costs, 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.

1.1 Literature

The Gittins index representation of optimal policies in discounted bandit problems with independent arms is not robust to perturbations of the standard model such as correlated arms, non-geometric discounting and the simultaneous pulling of multiple arms. In particular, Banks and Sundaram (1994) show that index policies are not optimal in the presence of switching costs between arms.[4] 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, are incurred whenever an inactive arm is not pulled in experimentation.[5] Nevertheless, optimal experimentation policies in bandit problems with maintenance costs fail to admit a Gittins index representation for the reason found by Banks and Sundaram (1994):

[3]This is reminiscent of the literature on escalation and the sunk cost fallacy. See Staw (1981), Staw and Ross (1987) and Garland (1990). McAfee et al. (2010) have expressed misgivings about interpreting such phenomena solely from a non-rational perspective.

[4]General characterisations of optimal experimentation policies with switching costs are not known. For details, see Jun (2004). An exception is Bergemann and Välimäki (2001), who exploit results of Banks and Sundaram (1992b) on bandits with a countably infinite set of ex ante identical arms.

[5]See the Conclusion for more discussion of the relationship between switching and maintenance costs.

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, breaking the independence required for an index characterisation.[6]

The exponential bandit framework of Keller et al. (2005) has proved useful in applications due to its tractability.[7] 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 (2010) generalise the model to Poisson bandits that allow for bad arms to also generate successes. Klein and Rady (2011) allow for negative correlation in the types of two experimenters' risky arms. Strulovici (2010) applies the model in a voting framework. In an earlier contribution, Bergemann and Hege (1998) introduce a discrete-time version of the model to study the moral hazard problem arising between bankers (principals) and venture capitalists (experimenters). In this vein, Bonatti and Hörner (2011), Hörner and Samuelson (2013) and Klein (2011) focus on the provision of incentives to experimenting agents.

Bonatti and Hörner (2011) derive another version of the stay-with-the-winner rule when agents can experiment with multiple disjunctive projects, i.e., when, as in my model, project successes are perfect substitutes. They also uncover a go-with-the-loser rule when projects are conjunctive, i.e., when successes on both projects are perfect complements. 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. Also, in Section 6, I show that going with the loser is still optimal if successes on various projects can be accumulated. Hence, my results are due to the experimenter's incentive to economise on maintenance costs by culling losing projects, and not to the perfect substitutability of project successes.

1.2 Example

The following simple 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, and can irreversibly discard any project at any time. Projects are risky in that the payoffs they deliver are unknown. A good project delivers a one-time payoff of 1, termed a success, with probability $\lambda > 0$ in any trial. A bad project never delivers a success. Assume that experimentation ends once a single trial is successful. Direct experimentation costs are $c > 0$ per trial, maintenance costs for a maintained but inactive project are $\gamma \geq 0$ per period and there is no discounting.

A project's current state is characterised by the experimenter's belief that it is good and repeated failures make the experimenter more pessimistic about the project. Let p^i_J be the

[6] Furthermore, 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.

[7] An earlier literature introduced learning about the arrival rate of a Poisson process to model R&D races. See Choi (1991) and Malueg and Tsutsui (1997).

probability that project J is good given that it has failed i trials, with $i \in \{0, 1, 2\}$. By Bayes' rule, $p_J^i = {p^{i-1}(1-\lambda)}/{(1-p^{i-1}\lambda)}$ for $i = \{1, 2\}$, and initial beliefs (p_A^0, p_B^0) are given. Assume that $p_A^2 \lambda > c$, so that in the absence of project B the experimenter would put project A to trial in all three periods, which also ensures that the experimenter never finds it optimal to discard both projects. Assume further 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 $\gamma = 0$, the optimal experimentation sequence is the stay-with-the-winner sequence AAB which allocates the first two trials to project A and the final trial to project B.

When $\gamma > 0$, the optimal experimentation sequence must be one of $AA|_A B$, $|_B AAA$ or $B|_B AA$, where $|_J$ represents the discarding of project J. In other words, either the experimenter sticks with the stay-with-the-winner rule, discards the losing project B prior to the first trial or she gives project B an early chance to succeed and discards it following a single failure. This follows since (a) if a project is discarded prior to the first trial, it must be project B, (b) if a project is discarded after a single trial, it must be project B irrespective of the project first put to trial, and among such sequences $B|_B AA$ yields the highest probability of a success, while (c) any experimentation sequence that maintains both projects in the first two trials must discard a project before the final trial, and among such sequences $AA|_A B$ yields the highest probability of a success. 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

$$
\begin{aligned}
v(AA|_A B; p_A^0, p_B^0) &= p_A^0\lambda + (1 - p_A^0\lambda)p_A^1\lambda + (1 - p_A^0\lambda)(1 - p_A^1\lambda)p_B^0\lambda \\
&\quad - \Big[(c+\gamma) + (1 - p_A^0\lambda)(c+\gamma) + (1 - p_A^0\lambda)(1 - p_A^1\lambda)c\Big], \\
v(B|_B AA; p_A^0, p_B^0) &= p_B^0\lambda + (1 - p_B^0\lambda)p_A^0\lambda + (1 - p_B^0\lambda)(1 - p_A^0\lambda)p_A^1\lambda \\
&\quad - \Big[(c+\gamma) + (1 - p_B^0\lambda)c + (1 - p_B^0\lambda)(1 - p_A^0\lambda)c\Big], \\
v(|_B AAA; p_A^0, p_B^0) &= p_A^0\lambda + (1 - p_A^0\lambda)p_A^1\lambda + (1 - p_A^0\lambda)(1 - p_A^1\lambda)p_A^2\lambda \\
&\quad - \Big[c + (1 - p_A^0\lambda)c + (1 - p_A^0\lambda)(1 - p_A^1\lambda)c\Big].
\end{aligned}
$$

The benefits of the go-with-the-loser sequence $B|_B AA$ are that (a) (relative to sequence $|_B AAA$) the value of project B gets exploited and the decision to discard B is better informed while (b) (relative to sequence $AA|_A B$) saving on maintenance costs. However, to the experimentation sequence $B|_B AA$ are associated both (a) the maintenance cost (relative to $|_B AAA$) and (b) 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 $p_B^0 \in [p_A^2, p_A^1]$, 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 $p_B^0 = p_A^2$.

That is, when $p_B^0 = p_A^2$, 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 λ, c, γ 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 $\lambda = 2/5$, $c = 6/100$, $\gamma = 3/200$ and $p_A^0 = 45/100$. Then it can be computed that $p_A^1 \approx 0.33$ and $p_A^2 \approx .23$, while $\underline{p} \approx 0.29$ and $\overline{p} \approx 0.32$. This is depicted in Figure 1.

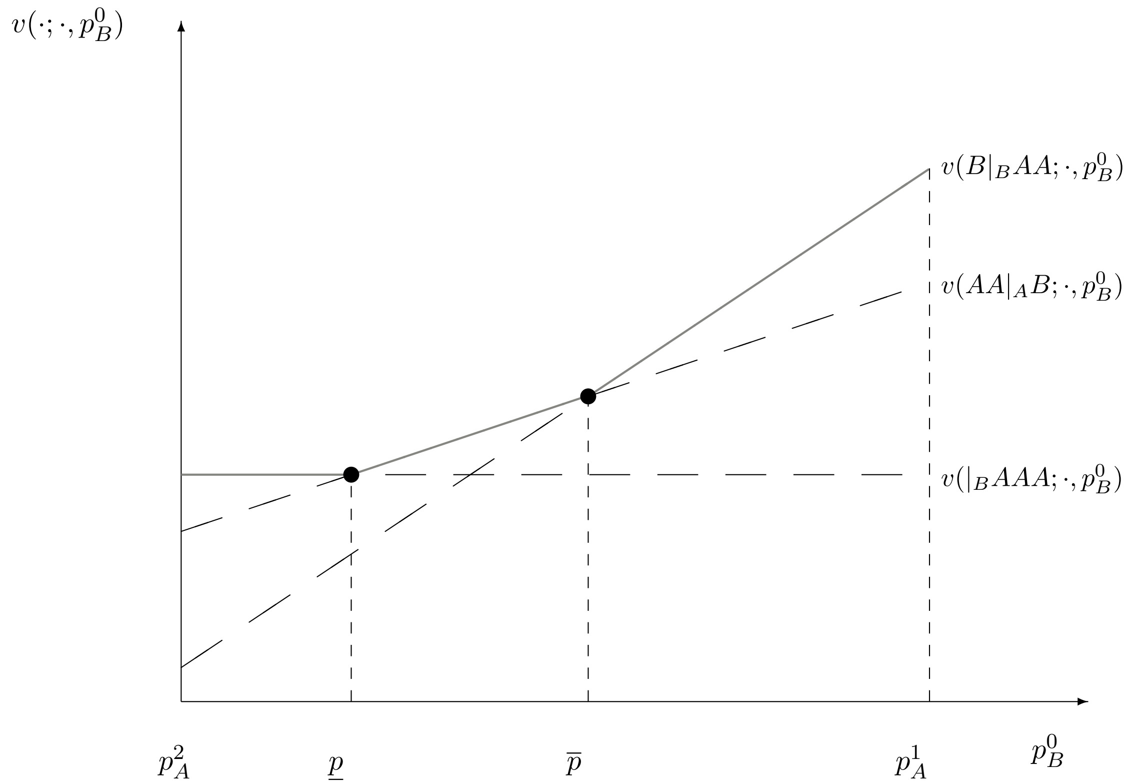

Figure 1: *Optimal experimentation in a three-period example.*

2 Model

I consider a continuous time three-armed exponential bandit problem with two risky arms, A and B, and a safe arm Q.[8] Risky arms are henceforth referred to as projects, and the safe arm as the option to quit experimentation. Running trials consists of experimenting with a risky project for some time interval $[t, t+dt)$. Trials yield either successes or failures. A risky project is either good or bad. A good project that is put to trial continuously in time interval $[t, t+dt)$ succeeds with probability λdt for some $\lambda > 0$, while a bad project fails with probability 1. The types of risky projects A and B are drawn independently, where (p_A^0, p_B^0) are the initial beliefs of the experimenter about the types of projects A and B, with p_J^0 the probability that project $J \in \{A, B\}$ is good. Without loss of generality, I assume that $p_A^0 \geq p_B^0$.

A risk-neutral experimenter collects the benefits of successes on risky projects and bears all experimentation costs. Its objective is to maximise its expected total discounted payoff. A success on either risky project yields a lump-sum payment of 1 and ends the experimentation process.[9] Quitting experimentation, which for simplicity I assume is irreversible, yields a payoff of 0. Putting risky project J to trial in time interval $[t, t+dt)$ entails direct experimentation cost cdt. I assume that $\lambda > c$, as otherwise the experimenter strictly prefers to quit experimentation at time $t = 0$. I introduce explicit costs to maintaining inactive risky projects. That is, a risky project that is maintained but not involved in trials in time interval $[t, t+dt)$ imposes cost γdt. The experimenter can irreversibly discard risky projects without cost and quitting experimentation also entails no costs. The experimenter discounts future payoffs at rate $r > 0$.

If at time t the experimenter has not quit experimentation, then in time interval $[t, t+dt)$ she receives a mass dt of trials to devote to any maintained projects. If both projects are maintained, the experimenter chooses the fraction αdt of trials to allocate to project A, where $\alpha \in [0, 1]$, with the remaining fraction $(1-\alpha)dt$ being allocated to project B. Since quitting experimentation is irreversible, if a single project is maintained, we have that $\alpha \in \{0, 1\}$. Note that when the experimentation occurs on both risky projects in time interval $[t, t+dt)$, the experimenter bears costs $(c+\gamma)dt$ irrespective of the allocation of trials received by either project. An allocation strategy $\{\alpha_t\}_{t\geq 0}$ is measurable with respect to the information available at time t. The experimenter also chooses stopping times t_A and t_B such that in the absence of a success project J is maintained at any time $t < t_J$, and discarded at t_J. The experimenter quits experimentation at time $\max\{t_A, t_B\}$.

At time t, beliefs (p_A^t, p_B^t) associated to maintained projects are a natural state variable capturing all payoff-relevant information about experimentation prior to t. Following Keller et al. (2005), if the experimenter allocates fraction αdt of trials to project A in time interval

[8] See the Conclusion for a discussion of results with more than two risky arms.
[9] See Section 6 for an extension to the case in which successes are accumulated across projects.

$[t, t+dt)$, the evolution of beliefs, by Bayes' rule, satisfies

$$dp_A^t = -\alpha\lambda p_A^t(1-p_A^t)dt, \text{ and}$$
$$dp_B^t = -(1-\alpha)\lambda p_B^t(1-p_B^t)dt.$$

It is without loss of generality to restrict attention to allocation strategies and discarding times generating belief processes $\{(p_A^t, p_B^t)\}_{t\geq 0}$ such that $p_A^t \geq p_B^t$ for all t. Since the expected arrival rates of successes depend only on beliefs, which decrease continuously when a project is put to trial and no success arrives, any strategy could be replaced by a payoff-equivalent strategy in the specified class through an appropriate relabelling of projects. Henceforth, project A will always be the 'winning' project, with project B the 'losing' project. Finally, it is without loss of generality for optimal experimentation to assume that $t_A \geq t_B$. On the one hand, it would never be optimal for the experimenter to discard project A if its belief was strictly higher than that of project B. On the other hand, if the beliefs of the two projects are identical and the experimenter finds it optimal to discard one of them, then she must be indifferent between discarding either project, so that in particular discarding project B is optimal.

Given beliefs (p_A^0, p_B^0), the experimenter chooses allocation strategy $\{\alpha_t\}_{t\geq 0}$ and discarding times (t_A, t_B) that maximise its expected total discounted payoff, given by

$$u(p_A^0, p_B^0) = \mathbb{E}\left[\int_0^{t_B} e^{-rt}[\alpha_t p_A^t\lambda + (1-\alpha_t)p_B^t\lambda - (c+\gamma)]dt + \int_{t_B}^{t_A} e^{-rt}[p_A^t\lambda - c]dt\right],$$

where the expectation is taken with respect to $\{\alpha_t\}_{t\geq 0}$ and $\{(p_A^t, p_B^t)\}_{t\geq 0}$.

3 The Bellman Equation

Once project B has been discarded, the experimenter faces a simple stopping problem with project A, and Keller et al. (2005) derive tractable expressions for its optimal payoff. Specifically, fix belief p_A and suppose that project B has been discarded, then the experimenter's value function must satisfy the Bellman equation

$$u(p_A) = \max\{0, [p_A\lambda - c]dt + e^{-rdt}\mathbb{E}[u(p_A + dp_A)|p_A]\}. \tag{1}$$

The first term in the brackets of (1) corresponds to quitting experimentation, and the second term corresponds to putting project A to trial. If project A is put to trial in a time interval of length dt, it succeeds with probability $p_A\lambda dt$, and the payoff to success is 1. The project fails with complementary probability, in which case the experimenter' payoff is $u(p_A) + u'(p_A)dp_A$, which is equal to $u(p_A) - \lambda p_A(1-p_A)u'(p_A)dt$. Using $1 - rdt$ as an approximation to e^{-rdt} and cancelling dominated terms, it follows that on any open set of beliefs p_A at which continuing experimentation at p_A is uniquely optimal, u satisfies the differential equation

$$u(p_A) = \frac{p_A\lambda}{r}\left[1 - u(p_A) - (1-p_A)u'(p_A)\right] - \frac{c}{r}.$$

Its general solution is

$$u(p_A) = Cz(p_A) + p_A \frac{\lambda - c}{r + \lambda} - (1 - p_A)\frac{c}{r}, \tag{2}$$

where C is a constant of integration and $z(p_A) = ({(1-p_A)}/{p_A})^{\frac{r}{\lambda}} (1 - p_A)$. The setup here is slightly different than in Keller et al. (2005), but the expression (2) admits a similar interpretation. The term $p_A {(\lambda-c)}/{(r+\lambda)} - (1-p_A){c}/{r}$ is the payoff to risky project A in the absence of the ability to quit experimentation, while the term $Cz(p_A)$ captures the option value of quitting experimentation. Imposing the value-matching and smooth-pasting conditions $u(p_A^*) = 0$ and $u'(p_A^*) = 0$ yields the quitting belief $p_A^* = {c}/{\lambda}$ and the constant of integration $C_D = ({c}/{(\lambda-c)})^{\frac{r}{\lambda}} {\lambda c}/{r(r+\lambda)}$. The following result, whose proof obtains through a verification argument in the Appendix, characterises optimal experimentation when project B has been discarded.

Lemma 1. *Fix belief p_A and suppose that project B has been discarded. If $p_A \leq {c}/{\lambda}$, then it is optimal for the experimenter to discard project A. If $p_A > {c}/{\lambda}$, then it is optimal for the experimenter to put project A to trial. The experimenter's resulting payoff is*

$$u_D(p_A) = \max\left\{C_D z(p_A) + p_A \frac{\lambda - c}{r + \lambda} - (1 - p_A)\frac{c}{r}, 0\right\}.$$

By manipulations mimicking those above, it follows that given any beliefs (p_A, p_B) at which no project has been discarded, the experimenter's value function must satisfy the Bellman equation[10]

$$u(p_A, p_B) = \max\left\{\max_{\alpha\in[0,1]}\left\{\alpha V_A[u](p_A, p_B) + (1-\alpha)V_B[u](p_A, p_B) - \frac{c+\gamma}{r}\right\}, u_D(p_A)\right\},$$

where

$$V_A[u](p_A, p_B) = \frac{p_A\lambda}{r}\left[1 - u(p_A, p_B) - (1 - p_A)\frac{\partial}{\partial p_A}u(p_A, p_B)\right], \text{ and}$$
$$V_B[u](p_A, p_B) = \frac{p_B\lambda}{r}\left[1 - u(p_A, p_B) - (1 - p_B)\frac{\partial}{\partial p_B}u(p_A, p_B)\right].$$

As in Keller et al. (2005), $V_J[u](p_A, p_B)$ captures the flow value of information generated by trials on project $J \in \{A, B\}$. In the Bellman equation, this is set against the flow costs of trials (the direct cost of project J and the maintenance cost of the other project). As the overall flow value of information is linear in α, the Bellman equation can be further simplified to

$$u(p_A, p_B) = \max\left\{V_A[u](p_A, p_B) - \frac{c+\gamma}{r}, V_B[u](p_A, p_B) - \frac{c+\gamma}{r}, u_D(p_A)\right\}. \tag{3}$$

[10]Note that given an allocation of trials $\alpha \in [0, 1]$ to project A that is constant in an interval of length dt, the probability of a single success on either project in this interval is $\alpha p_A\lambda dt + (1 - \alpha)p_B\lambda dt$.

4 Optimal Experimentation without Maintenance Costs

To highlight the impact of maintenance costs on optimal experimentation, a useful benchmark is the model without maintenance costs, that is, in which $\gamma = 0$. Note that this problem is a standard three-armed bandit problem with direct experimentation flow cost c, and the next result shows that optimal experimentation involves staying with the winner.

Proposition 1. *Suppose that $\gamma = 0$, fix beliefs (p_A, p_B) with $p_A \geq p_B$, and assume that no project has been discarded so far. If $p_A \leq c/\lambda$, then it is optimal for the experimenter to quit experimentation. If $p_A > c/\lambda$ and $p_A > p_B$, then it is optimal for the experimenter to put project A to trial. If $p_A = p_B > c/\lambda$, then it is optimal for the experimenter to share trials equally between both projects ($\alpha = 1/2$).*

To proceed quickly to the more interesting case in which $\gamma > 0$, I leave all details concerning the construction of the value function when $\gamma = 0$, as well as the verification that it satisfies the Bellman equation (3), to the Appendix. In the absence of maintenance costs, no project is discarded while experimentation is ongoing and the project with the highest belief is put to trial. This mimics the Gittins index representation of the optimal experimentation policy if a project's belief is taken to be the index.[11] Continuing experimentation is optimal as long as one project's belief is above the cutoff c/λ, which, since $p_A \geq p_B$, holds as long as $p_A > c/\lambda$. Figure 2 illustrates belief dynamics consistent with optimal experimentation. Given belief (p'_A, p'_B) with $p'_A > c/\lambda > p'_B$, only project A is ever put to trial, until $p_A = c/\lambda$, at which both projects are discarded. From initial belief (p''_A, p''_B) with $p''_A > p''_B > c/\lambda$, it is optimal to experiment with project A, followed by shared experimentation, until the beliefs of both projects reach c/λ.

5 Optimal Experimentation with Maintenance Costs

I characterise the optimal experimentation policy with maintenance costs in a number of steps. The main results of the paper are detailed in Section 5.1, where I describe the optimal discarding decision for the losing project B, conditional on putting it to trial ahead of the winning project A. As a byproduct of these results, I derive a simple necessary and sufficient condition for going with the loser to feature in the optimal experimentation policy. I also show that going with the loser must involve a simple culling rule: project B is given priority only if the experimenter expects to discard it before returning to project A. In Section 5.2, I tackle the experimenter's choice between staying with the winner, which may involve shared experimentation, and putting the losing project through a culling phase. While characterising this decision fully is not tractable, I build on the results of Section 5.1 to provide sufficient conditions for the experimenter to put the

[11] Exponential bandit problems belong to a class of problems for which 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 to a class of two-type symmetric bandit problems by Banks and Sundaram (1992a).

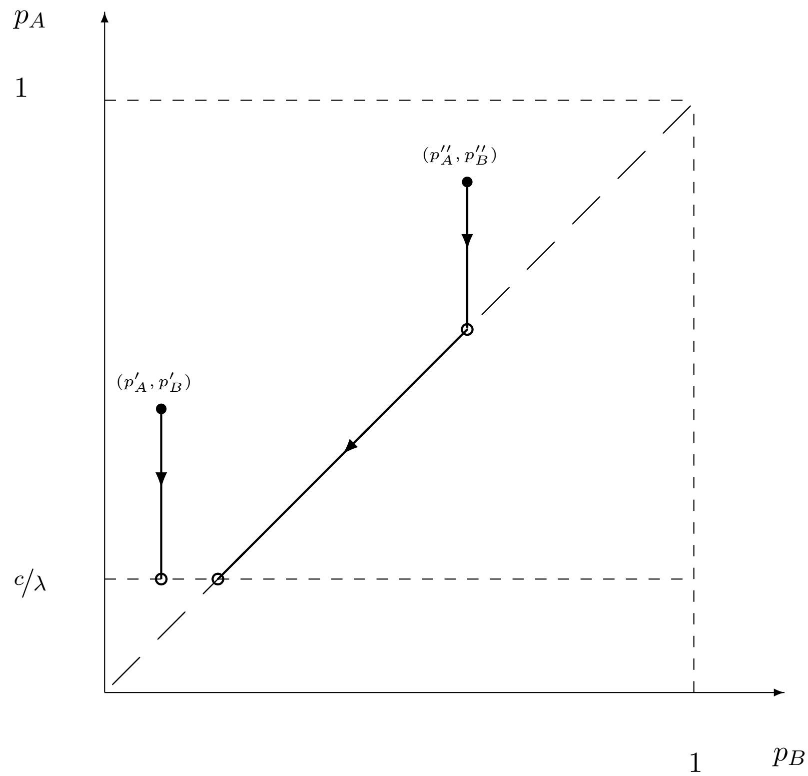

Figure 2: *Optimal experimentation without maintenance costs.*

winning project to trial in an open set of beliefs at which both projects are maintained. Proceeding in this way also illustrates the surprisingly complex switching of trials between projects under the optimal experimentation policy under maintenance costs. In Section 5.3, I discuss how to use the results of the two previous sections to complete the description of the optimal policy.

5.1 Discarding vs. Going with the Loser

If (p_A, p_B) lie in an open set of beliefs in which it is strictly optimal to put project B to trial, the experimenter's payoff satisfies

$$u(p_A, p_B) = V_B[u](p_A, p_B) - \frac{c+\gamma}{r} > \max\left\{V_A[u](p_A, p_B) - \frac{c+\gamma}{r}, u_D(p_A)\right\}.$$

The ordinary differential equation appearing here has the general solution

$$u(p_A, p_B) = Cz(p_B) + p_B \frac{\lambda - [c+\gamma]}{r+\lambda} - (1-p_B)\frac{c+\gamma}{r},$$

where C is a constant of integration. If project B is discarded at beliefs (p_A, p_B^*) with $p_B^* < p_B$, then the experimenter's payoff will satisfy the relevant value-matching and smooth-pasting conditions

$$u(p_A, p_B^*) = u_D(p_A), \text{ and} \tag{4}$$

$$\frac{\partial}{\partial p_B} u(p_A, p_B^*) = \frac{\partial}{\partial p_B} u_D(p_A) = 0, \tag{5}$$

which, when imposed on the above differential equation, yields

$$p_B^* = \frac{r u_D(p_A) + c + \gamma}{\lambda[1 - u_D(p_A)]} \equiv f(p_A), \text{ and} \tag{6}$$

$$C_L(p_A) = \frac{u_D(p_A) - \left[f(p_A)\frac{\lambda-[c+\gamma]}{r+\lambda} - (1-f(p_A))\frac{c+\gamma}{r} \right]}{z(f(p_A))}.$$

Finally, define the payoff function

$$u_L(p_A, p_B) = C_L(p_A)z(p_B) + p_B \frac{\lambda - [c+\gamma]}{r+\lambda} - (1-p_B)\frac{c+\gamma}{r}.$$

Note that for a fixed $p_A \in [c/\lambda, 1]$, there exists some belief p_B such that $p_B > f(p_A)$ only if $p_A > f(p_A)$. Also, since $u_D''(p_A) > 0$ for all $p_A \in [c/\lambda, 1]$, we have that $f''(p_A) > 0$, and hence it follows that there exist beliefs (p_A, p_B) with $p_A \geq p_B > f(p_A)$ if and only if

$$\begin{aligned} F(\lambda, r, c, \gamma) &\equiv \max_{p_A \in [c/\lambda, 1]} p_A - f(p_A) \\ &> 0, \end{aligned}$$

in which case there exist $\underline{p} < \overline{p}$ such that $\underline{p} - f(\underline{p}) = \overline{p} - f(\overline{p}) = 0$ and $p_A > f(p_A)$ for all $p_A \in (\underline{p}, \overline{p})$. Since $\partial/\partial\gamma f(p_A) > 0$, the envelope theorem implies that $\partial/\partial\gamma F(\lambda, r, c, \gamma) < 0$, yielding the following claim: *given (λ, r, c), there exists a cutoff maintenance cost $\tilde{\gamma}$ such that $p_A \geq p_B > f(p_A)$ for some beliefs (p_A, p_B) if and only if $\gamma < \tilde{\gamma}$. Furthermore, since $F(\lambda, r, c, 0) > 0$, it follows that* $\tilde{\gamma} > 0$. Given beliefs (p_A, p_B) with $p_A \geq p_B$, the belief $f(p_A) < p_B$ dictates when the experimenter should optimally discard project B, conditional on putting it to trial. The previous claim ensures that there is a non-negligible set of maintenance costs (i.e., $\gamma < \tilde{\gamma}$) for which some beliefs (p_A, p_B) admit candidate discarding beliefs $(p_A, f(p_A))$. As illustrated in Figure 3, this condition requires that the graph of f has a section that lies above the 45-degree line. The following result, whose proof follows from a verification argument in the Appendix, shows that for $\gamma \geq \tilde{\gamma}$, having multiple projects available yields no additional value.

Proposition 2. *Fix beliefs* (p_A, p_B) *with* $p_A \geq p_B$*, and assume that no project has been discarded so far. If* $\gamma \geq \tilde{\gamma}$*, then it is optimal to discard project* B*.*

The graph of f when $\gamma \in (0, \tilde{\gamma})$ can be characterised further (refer again to Figure 3 for an illustration). First, since $u_D(1) = {}^{(\lambda-c)}/_{(r+\lambda)}$, it follows that $1 - f(1) = -{}^{\gamma[r+\lambda]}/_{\lambda[r+c]} < 0$: the experimenter cannot be given the incentive to experiment with project B when project A is almost sure to be good. Second, since $u_D({}^{c}/_{\lambda}) = 0$, it follows that ${}^{c}/_{\lambda} - f({}^{c}/_{\lambda}) = -{}^{\gamma}/_{\lambda} < 0$: the experimenter does not have the incentive to experiment with project B when the beliefs of both projects are close to the discarding cutoff in the absence of maintenance costs. Third, that $u'_D(p_A) > 0$ for all $p_A \in ({}^{c}/_{\lambda}, 1]$ implies that $f'(p_A) > 0$: the experimenter discards project B earlier when she is more optimistic about the quality of project A. Fourth, as noted above, $f(p_A)$ is increasing in the maintenance cost γ: conditional on putting the losing project to trial, the experimenter discards it earlier when the cost of maintaining the winning project is higher. The comparative statics of the discarding belief $f(p_A)$ in the discount rate r, the success rate for good projects λ and the direct experimentation cost c are ambiguous in general: intuitively, the payoffs to retaining or discarding project B are both decreasing in r and c and increasing in λ.

The preceding results do not address whether putting project B to trial at beliefs (p_A, p_B) with $p_A > p_B$ is ever optimal. I now show that the condition that $\gamma \in (0, \tilde{\gamma})$ is not only necessary but also sufficient for going with the loser to be optimal at some beliefs. To start, recall that if beliefs (p_A, p_B) with $p_A \geq p_B > f(p_A)$ are such that going with the loser is strictly optimal, we have that $V_B[u_L](p_A, p_B) - {}^{(c+\gamma)}/_{r} > V_A[u_L](p_A, p_B) - {}^{(c+\gamma)}/_{r}$, which, when it is also the case that $p_A = p_B$, reduces to ${}^{\partial}/_{\partial p_B} u_L(p_B, p_B) < {}^{\partial}/_{\partial p_A} u_L(p_B, p_B)$. Next, define the belief

$$\check{p} = \max\left\{p \in [\underline{p}, \overline{p}] : \frac{\partial}{\partial p_A} u_L(p', p') \geq \frac{\partial}{\partial p_B} u_L(p', p') \text{ for all } p' \in [\underline{p}, p]\right\},$$

and note that, since (5) ensures that

$$\begin{aligned} \frac{\partial}{\partial p_A} u_L(\underline{p}, \underline{p}) &= u'_D(\underline{p}) \\ &> 0 \\ &= \frac{\partial}{\partial p_B} u_L(\underline{p}, \underline{p}), \end{aligned}$$

the continuous differentiability of u_L implies that $\check{p} > \underline{p}$. Finally, $\check{p} = \overline{p}$ if and only if ${}^{\partial}/_{\partial p_A} u_L(p, p) \geq {}^{\partial}/_{\partial p_B} u_L(p, p)$ for all $p \in [\underline{p}, \overline{p}]$. The next result uses the cutoff belief $\check{p}$ to identify an open set of beliefs at which putting project B to trial when $p_A > p_B$ is optimal, a stark reversal of the optimal experimentation policy relative to the case without maintenance costs.

Proposition 3. *Suppose that* $\gamma \in (0, \tilde{\gamma})$*, fix beliefs* (p_A, p_B) *with* $\check{p} \geq p_A \geq p_B$*, and assume that no project has been discarded so far.*

1. *If $p_A > {}^c\!/\!_\lambda$ and $p_B > f(p_A)$, then it is optimal for the experimenter to maintain both projects and put project B to trial.*

2. *If $p_A > {}^c\!/\!_\lambda$ and $p_B \leq f(p_A)$, then it is optimal for the experimenter to discard project B and put project A to trial.*

3. *If $p_A \leq {}^c\!/\!_\lambda$, then it is optimal for the experimenter to discard both projects.*

The result of Proposition 3, whose proof follows from a verification argument in the Appendix, is illustrated in Figure 3. Going with the loser is optimal at beliefs (p'_A, p'_B): in the absence of a success, the experimenter's belief that project B is good drops to $f(p'_A)$, after which project B is discarded and the experimenter puts project A to trial until its belief drops to ${}^c\!/\!_\lambda$. Beliefs (p'_A, p''_B) are such that $p''_B < f(p'_A)$, so that discarding project B and putting project A to trial is optimal.

An interesting feature of the optimal policy described in Proposition 3 is that putting project B to trial ahead of project A involves a simple culling rule: a losing project can be given priority only in the form of a 'last chance' to produce a success, with continued failure in this period of reprieve leading to the abandonment of the project. In fact, this is a general feature of the optimal experimentation policy: given beliefs (p_A, p_B) with $p_A > p_B$, any optimal policy that allocates some trials to project B can never allocate subsequent trials to project A without first discarding project B.[12] Intuitively, Proposition 1 ensures that, *conditional on maintaining both projects*, going with the winner is optimal, so that any period of experimentation that fails to follow this rule must end with the discarding of project B

5.2 Staying with the Winner vs. Going with the Loser

If $\check{p} = \overline{p}$, then putting project A to trial while maintaining project B is never optimal. In that case, Proposition 3 describes the optimal policy for all beliefs (p_A, p_B) with $\overline{p} \geq p_A \geq p_B$, while for all beliefs (p_A, p_B) with $p_A \geq p_B$ and $p_A > \overline{p}$, discarding project B is optimal.[13] Accordingly, for the remainder of this section, I assume that $\check{p} < \overline{p}$. As I show in Proposition 4 below, this condition is not only necessary, but also sufficient for staying with the winner to be optimal in an open set of beliefs in which both projects are maintained.

Staying with the winner at beliefs (p_A, p_B) can take two forms: either $p_A = p_B$ and trials are shared equally between both projects, or $p_A > p_B$ and project A is put to trial exclusively. To start, suppose that (p_B, p_B) with $p_B > \check{p}$ lie in an open set of beliefs in which it is strictly

[12] It can be shown independently that this property is necessary for optimality, although my results in Propositions 3 and 4 verify directly that u_L is the only payoff function satisfying the Bellman equation (3) for beliefs (p_A, p_B) with $p_A > p_B$ at which project B is put to trial.

[13] Verifying that the associated payoff function solves the Bellman equation (3) would follow from arguments in the proof of Proposition 3. In particular, note that given any $p_B \leq \overline{p}$, $u(\overline{p}, p_B) = u_D(\overline{p})$, where u is the value function from Proposition 3. Hence, the value-matching property holds at the boundary of the set of beliefs $\{(p_A, p_B) : p_A \geq p_B \geq \overline{p}\}$.

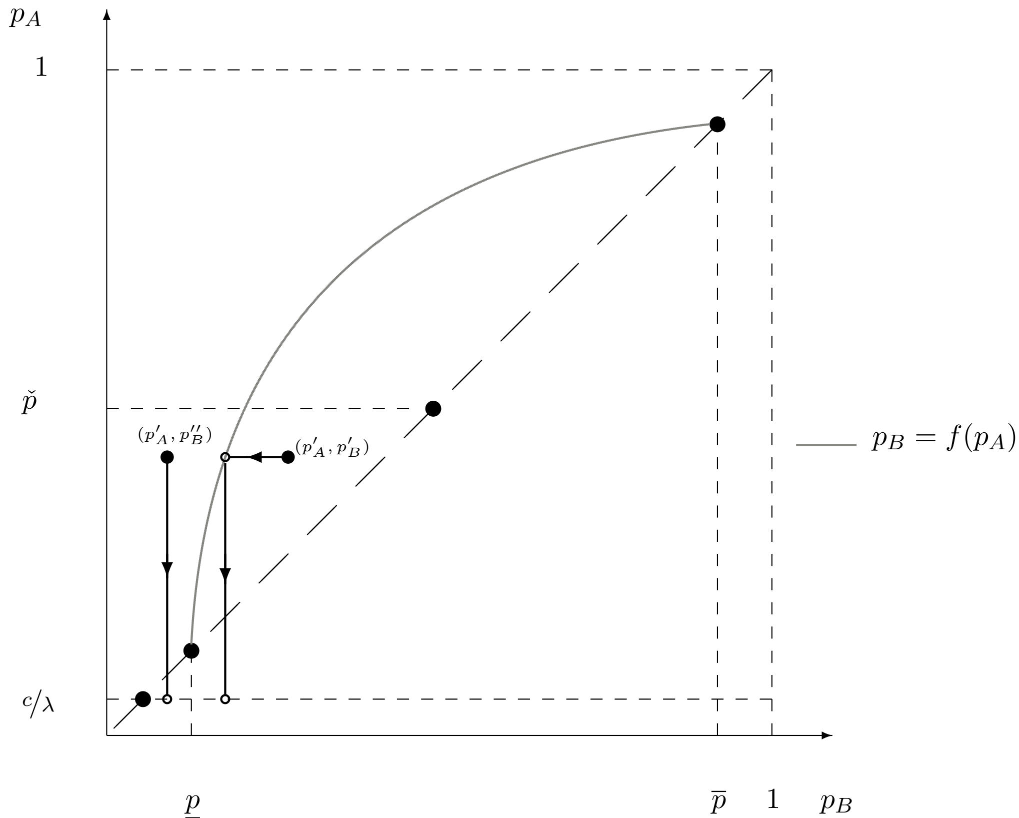

Figure 3: *The graph of f when $\gamma \in (0, \tilde{\gamma})$. Going with the loser is optimal at beliefs (p'_A, p'_B), while discarding project B is optimal at beliefs (p'_A, p''_B).*

optimal to share experimentation until the belief that each project is good drops to $\check{p}$ at which point, as is shown to be optimal in Proposition 3, the experimenter goes with the loser. Then the experimenter's payoff satisfies

$$u(p_B, p_B) = V_A[u](p_B, p_B) - \frac{c+\gamma}{r} = V_B[u](p_B, p_B) - \frac{c+\gamma}{r} > u_D(p_B).$$

The two differential equations appearing here can be combined to yield

$$u(p_B, p_B) = \frac{1}{2}\left[V_A[u](p_B, p_B) + V_B[u](p_B, p_B)\right] - \frac{c+\gamma}{r}.$$

If we let $u_E(p_B)$ denote the particular solution to this differential equation expressed solely as a function of the common belief p_B, we have that

$$u'_E(p_B) = \frac{\partial}{\partial p_A} u(p_B, p_B) + \frac{\partial}{\partial p_B} u(p_B, p_B),$$

and so it follows that

$$u_E(p_B) = C_E z(p_B)^2 + p_B^2 \frac{\lambda - [c+\gamma]}{r+\lambda} + 2p_B(1-p_B)\frac{\frac{\lambda}{2} - [c+\gamma]}{r+\frac{\lambda}{2}} - (1-p_B)^2 \frac{c+\gamma}{r},$$

where the constant of integration

$$C_E = \frac{u_L(\check{p},\check{p}) - \left[\check{p}^2 \frac{\lambda-[c+\gamma]}{r+\lambda} + 2\check{p}(1-\check{p})\frac{\frac{\lambda}{2}-[c+\gamma]}{r+\frac{\lambda}{2}} - (1-\check{p})^2 \frac{c+\gamma}{r}\right]}{z(\check{p})^2},$$

is obtained by imposing the value-matching condition

$$u_E(\check{p}) = u_L(\check{p},\check{p}). \tag{7}$$

The expression for u_E captures the fact that while the belief p_B is common to both projects, a success can arrive on either project, and that furthermore this belief decays at a slower rate than when a single project is put to trial, as each project receives only half the experimentation resources in any time interval.

Now suppose that (p_A, p_B) lie in an open set of beliefs in which it is strictly optimal to put project A to trial. The experimenter's payoff at (p_A, p_B) satisfies

$$u(p_A, p_B) = V_A[u](p_A, p_B) - \frac{c+\gamma}{r} > \max\left\{V_B[u](p_A, p_B) - \frac{c+\gamma}{r}, u_D(p_A)\right\}. \tag{8}$$

The ordinary differential equation appearing here has the general solution

$$u(p_A, p_B) = Cz(p_A) + p_A \frac{\lambda - [c+\gamma]}{r+\lambda} - (1-p_A)\frac{c+\gamma}{r},$$

where C is a constant of integration.

For beliefs (p_A, p_B) with $p_A \geq p_B \geq \check{p}$, let $u_W(p_A, p_B)$ be the particular solution to the differential equation in (8) where the constant of integration

$$C_W(p_B) = \frac{u_E(p_B) - \left[p_B \frac{\lambda-[c+\gamma]}{r+\lambda} - (1-p_B)\frac{c+\gamma}{r}\right]}{z(p_B)},$$

is obtained by imposing the value-matching condition

$$u_W(p_B, p_B) = u_E(p_B). \tag{9}$$

That is, for such beliefs, $u_W(p_A, p_B)$ describes the payoff from putting project A to trial until its belief drops to $p_B \geq \check{p}$, after which trials are shared until the common belief reaches $\check{p}$ and the experimenter, as is shown to be optimal in Proposition 3, goes with the loser.

Now fix beliefs (p_A, p_B) with $p_B \in [\underline{p}, \check{p})$. Proposition 3 implies that it cannot be optimal to put project A to trial until beliefs (p_B, p_B), since going with the loser is optimal at $(\check{p}, p_B)$. Instead, as I will verify in Proposition 4 below, conditional on putting project A to trial, the experimenter will go with the loser at beliefs (p_A^*, p_B) with $p_A^* > \check{p}$. To construct the payoff function associated to staying with the winner at (p_A, p_B), a first step is to identify this switching belief (p_A^*, p_B). To this end, note that if $V_A[u_L](p_A, p_B) > V_B[u_L](p_A, p_B)$, then the experimenter strictly prefers (a) putting project A to trial and delaying going with the loser for a small time interval dt at beliefs (p_A, p_B) to (b) going with the loser immediately, whereas if $V_A[u_L](p_A, p_B) < V_B[u_L](p_A, p_B)$, then the experimenter strictly prefers option (b) to option (a). At an optimal cutoff belief, it must be that $V_A[u_L](p_A^*, p_B) = V_B[u_L](p_A^*, p_B)$. In particular, note that $V_A[u_L](\check{p}, \check{p}) = V_B[u_L](\check{p}, \check{p})$, since $\partial/\partial p_A u_L(\check{p}, \check{p}) = \partial/\partial p_B u_L(\check{p}, \check{p})$.

Applying the implicit function theorem, we obtain a differentiable function g such that $g(\check{p}) = \check{p}$ and $V_A[u_L](g(p_B), p_B) = V_B[u_L](g(p_B), p_B)$ for all $p_B \leq \check{p}$; see Lemma 2 in the Appendix for details. This lemma also shows that g is decreasing on a domain $[\tilde{p}, \check{p}]$ with $\tilde{p} \in (\underline{p}, \check{p})$, and that furthermore the graphs of f and g never cross. This is illustrated in Figure 4. Finally, for beliefs (p_A, p_B) with $p_A \geq p_B$ and $p_B < \check{p}$ at which $g(p_B)$ is defined, let $u_W(p_A, p_B)$ be the particular solution to the differential equation in (8) where the constant of integration

$$C_W(p_B) = \frac{u_L(g(p_B), p_B) - \left[g(p_B)\frac{\lambda - [c+\gamma]}{r+\lambda} - (1 - g(p_B))\frac{c+\gamma}{r}\right]}{z(g(p_B))},$$

is obtained by imposing the value-matching condition

$$u_W(g(p_B), p_B) = u_L(g(p_B), p_B). \tag{10}$$

That is, for such beliefs, $u_W(p_A, p_B)$ describes the payoff from putting project A to trial until its belief drops to $g(p_B) > \check{p}$, after which the experimenter goes with the loser.

Define the function $h(p_B)$ such that, for beliefs $p_B \in [\check{p}, \hat{p}]$,

$$h(p_B) = \min\{p_A \geq p_B : u_W(p'_A, p_B) \geq \max\{u_L(p'_A, p_B), u_D(p'_A)\} \text{ for all } p'_A \in [p_B, p_A]\},$$

while for beliefs $p_B < \check{p}$ at which $g(p_B)$ is defined,

$$h(p_B) = \min\{p_A \geq g(p_B) : u_W(p'_A, p_B) \geq \max\{u_L(p'_A, p_B), u_D(p'_A)\} \text{ for all } p'_A \in [g(p_B), p_A]\}.$$

In words, $h(p_B)$ describes the first belief p_A for which, working backward in time starting from beliefs (p_B, p_B) when $p_B \geq \check{p}$ and from beliefs $(g(p_B), p_B)$ when $p_B < \check{p}$, the experimenter weakly prefers the best option from (a) going with the loser and (b) discarding project B and experimenting with project A only, to staying with the winner. Finally, define the belief

$$\hat{p} = \min\left\{p \geq \check{p} : \frac{\partial}{\partial p_A} u_L(p', p') \leq \frac{\partial}{\partial p_B} u_L(p', p') \text{ for all } p' \in [\check{p}, p]\right\}.$$

Note that $\check{p} < \overline{p}$ implies that $\check{p} < \hat{p} < \overline{p}$. First, (5) ensures that

$$\begin{aligned} \frac{\partial}{\partial p_A} u_L(\overline{p}, \overline{p}) &= u_D'(\overline{p}) \\ &> 0 \\ &= \frac{\partial}{\partial p_B} u_L(\overline{p}, \overline{p}), \end{aligned}$$

so that the continuous differentiability of u_L implies that $\hat{p} < \overline{p}$. Second, since $\check{p} < \overline{p}$ only if ${}^{\partial}/_{\partial p_A} u_L(p, p) < {}^{\partial}/_{\partial p_B} u_L(p, p)$ for all $p > \check{p}$ close to $\check{p}$, it follows that $\hat{p} > \check{p}$. As I show in Proposition 4 below, $(\hat{p}, \hat{p})$ are the highest beliefs at which shared experimentation until the common belief reaches $\check{p}$ is optimal. Figure 4 illustrates the graph of h, and Lemma 2 in the Appendix shows that h is increasing on $[\tilde{p}, \hat{p}]$, with $h(\tilde{p}) = g(\tilde{p})$, $h(\hat{p}) = \hat{p}$, $h(p_B) > g(p_B)$ for all $p_B \in (\tilde{p}, \check{p})$ and $h(p_B) > p_B$ for all $p_B \in [\check{p}, \hat{p})$. A final note, also illustrated in Figure 4, is that the graphs of f and h may cross. In Section 5.3 below, I discuss a numerical example in which this is indeed the case.

Define the stay-with-the-winner region

$$\mathcal{P}_W = \big\{(p_A, p_B) : p_B \in (\tilde{p}, \check{p}),\, g(p_B) < p_A < h(p_B)\big\} \\ \cup \big\{(p_A, p_B) : p_B \in [\check{p}, \hat{p}),\; p_B \leq p_A < h(p_B)\big\},$$

and the belief

$$\mathring{p} = \max\left\{p \in [\hat{p}, \overline{p}] : \frac{\partial}{\partial p_A} u_L(p', p') \geq \frac{\partial}{\partial p_B} u_L(p', p') \text{ for all } p' \in [\hat{p}, p]\right\}.$$

Together with Proposition 3, the following result characterises the optimal experimentation policy for beliefs (p_A, p_B) with $\mathring{p} > p_A \geq p_B$.

Proposition 4. *Suppose that $\gamma \in (0, \tilde{\gamma})$ and $\check{p} < \overline{p}$, fix beliefs (p_A, p_B) with $p_A \geq p_B$ and $p_A \in (\check{p}, \mathring{p})$, and assume that no project has been discarded so far.*

1. *If $(p_A, p_B) \in \mathcal{P}_W$, then it is optimal for the experimenter to maintain both projects, to put project A to trial when $p_A > p_B$, and to allocate experimentation equally between both projects when $p_A = p_B$.*

2. *If $(p_A, p_B) \notin \mathcal{P}_W$ and $p_B > f(p_A)$, then it is optimal for the experimenter to maintain both projects and put project B to trial.*

3. If $(p_A, p_B) \notin \mathcal{P}_W$ and $p_B \leq f(p_A)$, then it is optimal for the experimenter to discard project B and put project A to trial.

A downside of this characterisation is that it involves the graphs of g and h, along with their relationships with the graph of f, which are not tractable. However, irrespective of the global properties of these graphs, the condition $\check{p} < \overline{p}$ ensures the existence of an open set of beliefs from which the optimal experimentation policy with maintenance costs involves substantial switching back and forth between both projects, with patterns that differ strikingly from the optimal policy without maintenance costs. In particular, while Proposition 3 shows that going with the loser must be followed by the discarding of project B, Propositon 4 shows that staying with the winner must be followed by a culling phase for project B in which project A, while no longer put to trial, is nevertheless maintained. Figure 4 illustrates these rich experimentation dynamics. Staying with the winner is optimal at beliefs (p'_A, p'_B) and (p''_A, p''_B). From beliefs (p'_A, p'_B), the experimenter first puts project A to trial until, in the absence of a success, its belief drops to p'_B, after which experimentation is shared until the common belief reaches $\check{p}$, after which project B enters a culling phase. From beliefs (p''_A, p''_B), the experimenter puts project A to trial but, in the absence of a success, going with the loser is optimal at beliefs $(g(p''_B), p''_B)$, before the belief of project A reaches that of project B.

5.3 Completing the Characterisation of the Optimal Policy

In one case, the results of Propositions 3 and 4 are enough to characterise the optimal experimentation policy for all beliefs. Specifically, if $\mathring{p} = \overline{p}$, then discarding project B is optimal at all beliefs (p_A, p_B) with $p_A \geq p_B \geq \overline{p}$ at which both projects are maintained.[14] For example, if $\lambda = r = 1$, $c = 1/2$ and $\gamma = 1/100$, then it can be computed that $\underline{p} \approx 0.510$, $\overline{p} \approx 0.986$, $\check{p} \approx 0.582$, $\tilde{p} \approx 0.566$, $\hat{p} \approx 0.961$, and that $\mathring{p} = \overline{p}$.[15] In fact, other than depicting the case of $\mathring{p} < \overline{p}$, Figure 4 is drawn to reflect the properties of this numerical example. In particular, it can be verified that the graph of h lies above the graph of f for all p_B in a sub-interval of $[\tilde{p}, \hat{p}]$.

If instead $\mathring{p} < \overline{p}$, then the results of Propositions 3 and 4 can still be exploited to describe the optimal experimentation policy. Consider a belief $\check{p}' \in [\mathring{p}, \overline{p}]$ such that $\partial/\partial p_A u_L(\check{p}', \check{p}') = \partial/\partial p_B u_L(\check{p}', \check{p}')$ and going with the loser is optimal at $(\check{p}', \check{p}')$, along with a belief $\hat{p}' \in (\check{p}', \overline{p})$ such that $\partial/\partial p_A u_L(\hat{p}', \hat{p}') = \partial/\partial p_B u_L(\hat{p}', \hat{p}')$ and $\partial/\partial p_A u_L(p', p') < \partial/\partial p_B u_L(p', p')$ for all $p' \in (\check{p}', \hat{p}')$. Such beliefs must exist if $\mathring{p} < \overline{p}$ (e.g., $\check{p}' = \mathring{p}$ identifies one candidate). But then $\check{p}'$ can be substituted for $\check{p}$ and $\hat{p}'$ can be substituted for $\hat{p}$ in the results of Section 5.2, and a version of Proposition 4 then

[14] Verifying that the associated payoff function solves the Bellman equation (3) at these beliefs would follow from arguments contained in the proofs of Propositions 3 and 4. In that case, $f(\overline{p}) = h(\overline{p})$ and, given any $p_B < \overline{p}$, $\lim_{p_A \to \overline{p}} u(p_A, p_B) = u_D(\overline{p})$, where u is the value function from Proposition 4. Hence, the value-matching property holds at the boundary of the set of beliefs $\{(p_A, p_B) : p_A \geq p_B \geq \overline{p}\}$.

[15] Although this last property holds in all numerical examples I have computed, I cannot prove that it must be satisfied in general.

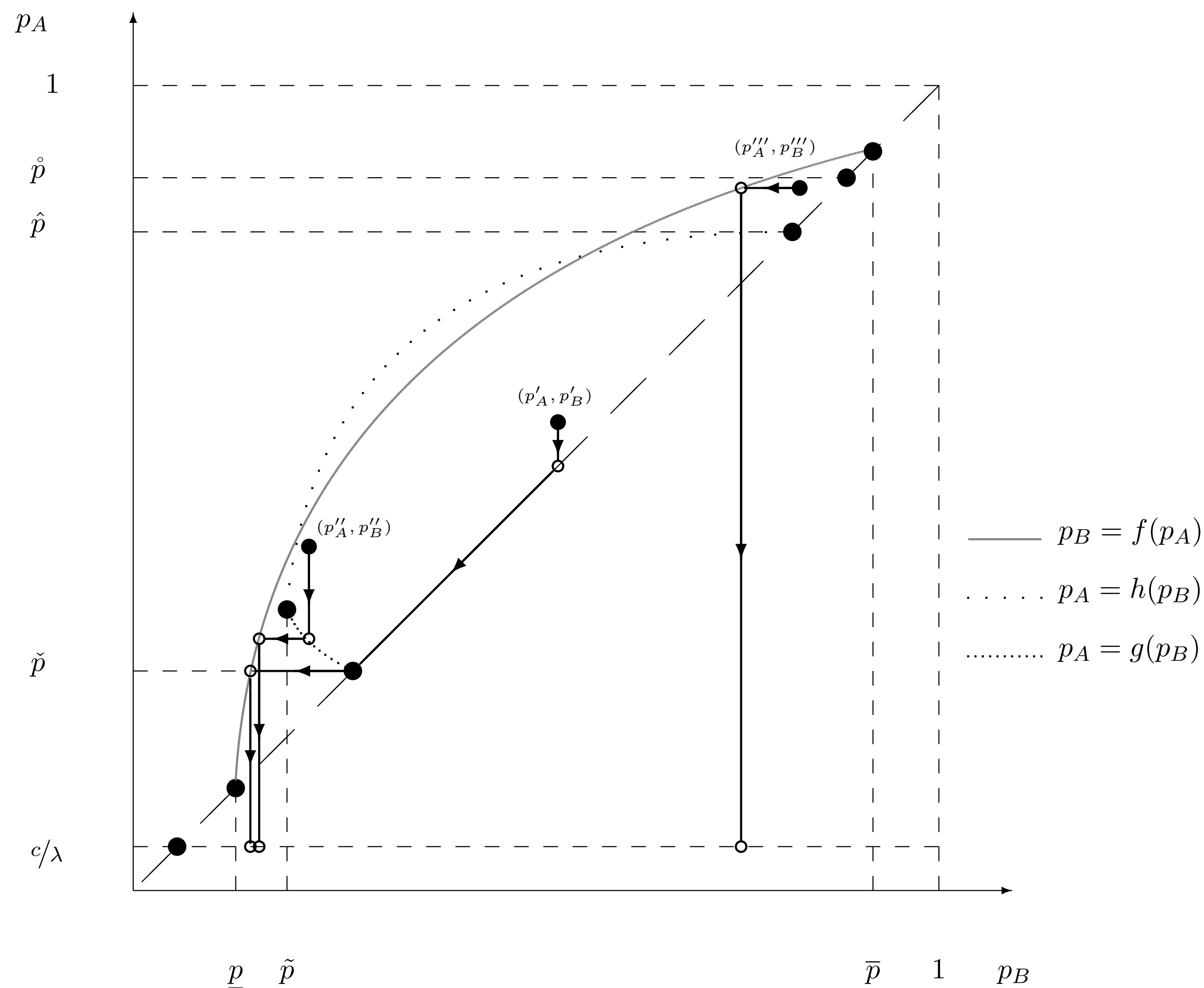

Figure 4: *The graphs of g and h when $\gamma \in (0, \tilde{\gamma})$ and $\check{p} < \overline{p}$. Staying with the winner is optimal at beliefs (p'_A, p'_B) and (p''_A, p''_B), while going with the loser is optimal at beliefs (p'''_A, p'''_B).*

describes the optimal experimentation policy for beliefs (p_A, p_B) with $p_A \geq p_B$ and $p_A \in (\check{p}', \mathring{p}')$, for some $\mathring{p}' \in (\check{p}', \underline{p}]$.

Since, as noted in Section 5.2, $\partial/\partial p_A u_L(p,p) > \partial/\partial p_B u_L(p,p)$ for all p close to $\overline{p}$, this extension of Proposition 4 implies that it must be optimal to go with the loser for beliefs (p_A, p_B) with $p_B > f(p_A)$ close to $(\overline{p}, \overline{p})$. Since, by Proposition 3, the same holds for beliefs close to $(\underline{p}, \underline{p})$, and furthermore by Proposition 4 staying with the winner is optimal for some beliefs (p,p) with $p \in (\underline{p}, \overline{p})$ (if $\check{p} < \underline{p}$), this highlights a noteworthy non-monotonicity in the optimal experimentation policy with maintenance costs. Intuitively, for both these sets of beliefs, it is as though the experimenter has decided not to maintain both projects in the long run. It remains that if a single project is to be maintained it should be project A. Yet, for beliefs (p_A, p_B) with $p_A \geq p_B > f(p_A)$, putting project B to trial yields strictly higher payoffs than discarding it immediately. There is an interesting difference in the experimenter's incentives in these two regions. For beliefs around $(\overline{p}, \overline{p})$, the experimenter maintains a single project in the long run because she is optimistic about the quality of both projects. Since both projects are likely to eventually yield a success, she puts the losing project B to trial in the hope of a quick success and then discards it if these trials fail and moves on to the winning project A. For beliefs around $(\underline{p}, \underline{p})$, the experimenter maintains a single project in the long run because she is pessimistic about the quality of both projects. Neither project is likely to yield a success, so that exploiting some of the residual value of project B and pinning its hopes on developing project A at a reduced cost is optimal. This difference in interpretation turns out to be important in Section 6 where I consider the case in which project successes can be accumulated. In that case, going with the loser when the experimenter is optimistic about both projects need not be optimal, but it remains optimal when the experimenter is pessimistic about both projects.

6 Additive Project Successes

I have assumed that the outcomes of the projects are perfect substitutes in that the experimenter searches for a single success on either project. An alternative assumption is that a successful project is retired but the experimenter obtains a payoff of 1 from all remaining projects that succeed. In this section, I do not fully characterise optimal experimentation with additive project successes, but instead show that if the set of beliefs at which maintaining both projects is optimal is nonempty, then it must contain beliefs at which putting project B to trial is optimal. Hence, the optimality of going with the loser is not due solely to the experimenter having an inelastic demand for successes. Rather, it is a key feature of minimising total project development costs in the presence of maintenance costs.

If (p_A, p_B) lie in an open set of beliefs in which it is strictly optimal to put project B to trial while project A is maintained, the experimenter's payoff satisfies the differential equation

$$u(p_A, p_B) = \frac{p_B\lambda[1 + u_D(p_A)] - [c + \gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1 - p_B)}{p_B\lambda + r}\frac{\partial}{\partial p_B}u(p_A, p_B).$$

The difference with similar expressions from previous sections is that when successes can be accumulated, the success of project B leads to continued experimentation with project A, so that the experimenter's payoff to a success on B is $1 + u_D(p_A)$. If project B is discarded at beliefs (p_A, p_B^*) with $p_B^* \leq p_B$, then value-matching and smooth-pasting conditions

$$u(p_A, p_B^*) = u_D(p_A) \text{ and}$$
$$\frac{\partial}{\partial p_B}u(p_A, p_B^*) = 0,$$

when imposed on the differential equation above, yield the following version of (6):

$$p_B^* = \frac{ru_D(p_A) + c + \gamma}{\lambda} \equiv \tilde{f}(p_A). \tag{11}$$

Since $1 - u_D(p_A) \in [0, 1]$, it follows that $\tilde{f}(p_A) \leq f(p_A)$. That is, when successes can be accumulated, the experimenter allocates additional trials to project B before discarding it. As in Section 5.1, there exists some belief p_B such that $p_B > \tilde{f}(p_A)$ if and only if, for some p_A, $p_A - \tilde{f}(p_A) > 0$. However, since $u_D(1) = {}^{(\lambda - c)}/_{(r+\lambda)}$, it need not be the case that $1 - \tilde{f}(1) < 0$, which, recalling Section 5.1, means that beliefs $(\overline{p}, \overline{p})$ such that $\overline{p} = \tilde{f}(\overline{p})$ need not exist. However, since $u_D({}^{c}/_{\lambda}) = 0$, it follows that ${}^{c}/_{\lambda} - \tilde{f}({}^{c}/_{\lambda}) = -{}^{\gamma}/_{\lambda} < 0$. That is, beliefs $(\underline{p}, \underline{p})$ such that $\underline{p} = \tilde{f}(\underline{p})$ with $\underline{p} > {}^{c}/_{\lambda}$ must exist.

Suppose, towards a contradiction, that staying with the winner is optimal in a neighbourhood of $(\underline{p}, \underline{p})$, so that the experimenter's payoff satisfies the differential equation

$$u(\underline{p}, \underline{p}) = \frac{\underline{p}\lambda[1 + u_D(\underline{p})] - [c + \gamma]}{\underline{p}\lambda + r} - \frac{\lambda\underline{p}(1 - \underline{p})}{\underline{p}\lambda + r}\frac{\partial}{\partial p_A}u(\underline{p}, \underline{p}).$$

Evaluated at $\underline{p} = \tilde{f}(\underline{p})$, (11) can be rewritten as

$$u_D(\underline{p}) = \frac{\underline{p}\lambda[1 + u_D(\underline{p})] - [c + \gamma]}{\underline{p}\lambda + r}.$$

Hence,

$$\begin{aligned} u(\underline{p}, \underline{p}) - u_D(\underline{p}) &= -\frac{\lambda\underline{p}(1 - \underline{p})}{\underline{p}\lambda + r}\frac{\partial}{\partial p_A}u(\underline{p}, \underline{p}) \\ &< 0, \end{aligned}$$

so that u cannot satisfy the Bellman equation of the problem with additive successes, yielding the desired contradiction. Intuitively, at beliefs close to $(\underline{p}, \underline{p})$, the experimenter strictly prefers discarding project B immediately to going with the winner, but for beliefs (p_A, p_B) with $p_B > f(p_A)$, she strictly prefers going with the loser to discarding project B immediately.

When successes can be accumulated, the experimenter has no incentive to go with the loser in order to maintain a single project in the long run when she is optimistic about both projects. This contrasts with the case of substitutable projects, in which a quick success on the winning project A when B is maintained kills the value of project B and sinks the expended maintenance costs. When successes are additive, a quick success on project A allows the experimenter to exploit project B's value promptly. However, when the experimenter is pessimistic about both projects, the likelihood of obtaining two successes is small and the experimenter has an incentive to economise on maintenance costs by going with the loser in order to quickly focus the brunt of its research efforts on the winning project.

7 Conclusion

The standard approach to experimentation assumes that keeping the option of researching projects that are not currently being put to trial 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 projects' option values and have important implications for optimal experimentation policies. In particular, my results provide a rationale for project development in which less promising alternatives receive priority over more promising ones.

While investigating the effects of maintenance costs on optimal experimentation in more general bandit settings is an interesting avenue for future work, the simple and tractable exponential bandit framework also allows for additional generalisations of my results. In particular, the culling rule for losing projects is key for characterising optimal experimentation with maintenance costs and it takes a more general form in the case in which the experimenter has more than two risky projects. When the experimenter has three risky projects ranked by their beliefs that they are good, it can be shown 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.[16] In other words, experimenting with a middle-ranked project grants a 'last chance' to all projects of a lower or equal rank.

A unified treatment of experimentation with both maintenance and switching costs also presents interesting possibilities. In an environment in which an experimenter can acquire new projects at a cost at any time, it is plausible that start-up costs for inactive projects are higher for those projects that have not been maintained. In that case, the experimenter would trade off (a) acquiring multiple projects early and paying to maintain inactive projects and (b) acquiring one

[16] This was shown in an earlier draft of the paper, and details are available upon request. The argument is easily extended to more than three projects.

project at a time and bearing switching costs whenever project development priorities change. A model along these lines could shed light on when a portfolio approach to experimentation is preferred to a more focused sequential approach. Experimentation dynamics would vary in these approaches, as an experimenter with a portfolio of projects purchases through maintenance costs the flexibility of re-prioritising between them easily, whereas an experimenter acquiring new projects as needed saves on maintenance costs by facing higher switching costs, and so should delay transitions between projects.

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 (2008). Bandit problems. *The New Palgrave Dictionary of Economics, 2nd ed.*.

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 (2011). Collaborating. *American Economic Review 101*(2), 632–663.

Choi, J. P. (1991). Dynamic R&D competition under "hazard rate" uncertainty. *The RAND Journal of Economics*, 596–610.

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.

Hörner, J. and L. Samuelson (2013). Incentives for experimenting agents. *Rand Journal of Economics 44*(4), 632–663.

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

Keller, G. and S. Rady (2010). Strategic experimentation with Poisson bandits. *Theoretical Economics 5*(2), 275–311.

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

Klein, N. (2011). The importance of being honest. *University of Montreal Working Papers*.

Klein, N. and S. Rady (2011). Negatively correlated bandits. *The Review of Economic Studies 78*(2), 693–732.

Malueg, D. A. and S. O. Tsutsui (1997). Dynamic R&D competition with learning. *The RAND Journal of Economics*, 751–772.

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. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica 78*(3), 933–971.

A Appendix

Proof of Lemma 1. Using the arguments from the text, the Bellman equation (1) can be rewritten as

$$u(p_A) = \max\left\{0, \frac{p_A\lambda}{r}\left[1 - u(p_A) - (1-p_A)u'(p_A)\right] - \frac{c}{r}\right\}. \tag{12}$$

Lemma 1 follows if I show that u_D satisfies (12). Suppose that $p_A \leq c/\lambda$. Hence, we have that $u_D(p_A) = 0$ and $u'_D(p_A) = 0$, and that

$$\begin{aligned}\frac{p_A\lambda}{r}\left[1 - u_D(p_A) - (1-p_A)u'_D(p_A)\right] - \frac{c}{r} &= \frac{p_A\lambda - c}{r}\\ &\leq 0,\end{aligned}$$

so that (12) is satisfied. Now suppose that $p_A > {}^c\!/\lambda$. We have that

$$\frac{p_A\lambda}{r}\left[1 - u_D(p_A) - (1-p_A)u_D'(p_A)\right] - \frac{c}{r} = u_D(p_A)$$
$$> 0,$$

so that (12) is satisfied. The inequality follows since $u_D({}^c\!/\lambda) = 0$ and $u_D'(p_A) > 0$ for $p_A > {}^c\!/\lambda$. □

Proof of Proposition 1. Fix beliefs (p_A, p_B) with $p_A \geq p_B$, and assume that no project has been discarded so far.

Step 1. A candidate value function, which I will show satisfies the Bellman equation (3), needs to be defined. Given $p \geq {}^c\!/\lambda$, define the function

$$\tilde{u}_E(p) = \tilde{C}_E z(p)^2 + p^2\frac{\lambda - c}{r+\lambda} + 2p(1-p)\frac{\frac{\lambda}{2} - c}{r + \frac{\lambda}{2}} - (1-p)^2\frac{c}{r},$$

where

$$\tilde{C}_E = \frac{c(r+\lambda-c)(\lambda-c)}{z(\frac{c}{\lambda})r(r+\lambda)(r+\frac{\lambda}{2})}.$$

Now given beliefs $p_A \geq p_B \geq {}^c\!/\lambda$, define the function

$$u_A(p_A, p_B) = C_A(p_B)z(p_A) + p_A\frac{\lambda - c}{r+\lambda} - (1-p_A)\frac{c}{r},$$

where

$$C_A(p_B) = \frac{\tilde{u}_E(p) - \left[p_B\frac{\lambda-c}{r+\lambda} - (1-p_B)\frac{c}{r}\right]}{z(p_B)}.$$

Intuitively, $\tilde{u}_E(p)$ describes the experimenter's payoff from beliefs (p,p) to sharing experimentation between both projects until beliefs drop to $({}^c\!/\lambda, {}^c\!/\lambda)$ when both projects are discarded, while $u_A(p_A, p_B)$ describes the experimenter's payoff from beliefs (p_A, p_B) to allocating all trials to project A until its belief drops to p_B, following which experimentation is shared until beliefs $({}^c\!/\lambda, {}^c\!/\lambda)$.[17] Hence, these payoffs are constructed to satisfy the value-matching and smooth-pasting conditions

$$\tilde{u}_E({}^c\!/\lambda) = 0,$$
$$\tilde{u}_E'({}^c\!/\lambda) = 0,$$
$$u_A(p_B, p_B) = \tilde{u}_E(p_B), \text{ and} \tag{13}$$
$$\frac{\partial}{\partial p_A}u_A(p_B, p_B) = \frac{1}{2}\tilde{u}_E'(p_B). \tag{14}$$

[17] An exhausive derivation of closely related payoff functions is presented in Section 5.2.

Finally, for all $p_A \geq p_B$, define

$$u(p_A, p_B) = \begin{cases} u_A(p_A, p_B) & \text{if } p_A > c/\lambda, \\ 0 & \text{if } p_A \leq c/\lambda. \end{cases}$$

Step 2. It needs to be verified that $u(p_A, p_B) = 0$ is a solution to the Bellman equation (3) when $c/\lambda \geq p_A \geq p_B$. Note that since $\frac{\partial}{\partial p_A} u(p_A, p_B) = 0$ and $\frac{\partial}{\partial p_B} u(p_A, p_B) = 0$, we have that

$$\begin{aligned} V_B(p_A, p_B) - \frac{c}{r} &= \frac{p_B \lambda - c}{r} \\ &\leq V_A(p_A, p_B) - \frac{c}{r} \\ &= \frac{p_A \lambda - c}{r} \\ &\leq 0 \\ &= u_D(p_A), \end{aligned}$$

so that (3) is satisfied.

Step 3. It needs to be verified that u_A is a solution to the Bellman equation (3) when $p_A > c/\lambda$. First, we have that

$$\begin{aligned} V_A(p_A, p_B) - \frac{c}{r} - u_D(p_A) &= z(p_A)[C_A(p_B) - C_D] \\ &\geq 0. \end{aligned}$$

Second, if $p_B \leq c/\lambda$, we can impose further that $u_A(p_A, p_B) = u_D(p_A)$. Hence, since $\frac{\partial}{\partial p_B} u_D(p_A) = 0$, we have that

$$\begin{aligned} V_A(p_A, p_B) - \frac{c}{r} - \left[V_B(p_A, p_B) - \frac{c}{r}\right] &= u_D(p_A) - \frac{1}{r}\left[p_B \lambda [1 - u_D(p_A)] - c\right] \\ &= u_D(p_A)\left[1 + \frac{p_B \lambda}{r}\right] - \left[\frac{p_B \lambda - c}{r}\right] \\ &> 0, \end{aligned}$$

where the first equality uses the fact that $V_A(p_A, p_B) - c/\lambda = u_D(p_A)$. Third, suppose that

$p_A > p_B > {}^c\!/\!_\lambda$. Note that

$$\begin{aligned}
&V_A(p_A,p_B) - \frac{c}{r} - \left[V_B(p_A,p_B) - \frac{c}{r}\right] \\
&\qquad = u_A(p_A,p_B) - \left[V_B(p_A,p_B) - \frac{c}{r}\right] \\
&\qquad = \frac{p_B\lambda + r}{r}\left[u_A(p_A,p_B) - \left[\frac{p_B\lambda - c}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r}\frac{\partial}{\partial p_B}u_A(p_A,p_B)\right]\right] \\
&\qquad = \frac{p_B\lambda + r}{r}\left[u_A(p_A,p_B) - \left[\frac{p_B\lambda - c}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r}C_A'(p_B)z(p_A)\right]\right] \\
&\qquad = \frac{p_B\lambda + r}{r}\left[u_A(p_A,p_B) - \frac{p_B\lambda - c}{p_B\lambda + r} - C_A(p_B)z(p_A) - \frac{z(p_A)}{z'(p_B)}\frac{\lambda[r+c]}{r[r+\lambda]}\right] \\
&\qquad = \frac{p_B\lambda + r}{r}\left[\left[p_A - \frac{z(p_A)}{z'(p_B)}\right]\frac{\lambda[r+c]}{r[r+\lambda]} - \frac{c}{\lambda} - \frac{p_B\lambda - c}{p_B\lambda + r}\right],
\end{aligned} \tag{15}$$

where the fourth equality follows since ${}^{z(p)}\!/\!_{z'(p)} = -{}^{\lambda p(1-p)}\!/\!_{(p\lambda+r)}$ and

$$\begin{aligned}
C_A'(p_B)z(p_B) &= \frac{\partial}{\partial p_B}u_A(p_B,p_B) \\
&= \frac{1}{2}\tilde{u}_E'(p_B) \\
&= \frac{\partial}{\partial p_A}u_A(p_B,p_B) \\
&= C_A(p_B)z'(p_B) + \frac{\lambda[r+c]}{r[r+\lambda]},
\end{aligned}$$

where the second equality follows from evaluating the total derivative of (13) and using (14). Using (15), we have that

$$\frac{\partial}{\partial p_A}\left[V_A(p_A,p_B) - \frac{c}{\lambda} - \left[V_B(p_A,p_B) - \frac{c}{\lambda}\right]\right] > 0,$$

if and only if ${}^{z'(p_A)}\!/\!_{z'(p_B)} < 1$, which holds since $p_A > p_B$ and it can be verified that, for all p, $z'(p) < 0$ and $z''(p) > 0$. Hence, since

$$V_A(p_B,p_B) - \frac{c}{\lambda} - \left[V_B(p_B,p_B) - \frac{c}{\lambda}\right] = 0,$$

it follows that, for $p_A > p_B$,

$$V_A(p_A,p_B) - \frac{c}{\lambda} - \left[V_B(p_A,p_B) - \frac{c}{\lambda}\right] > 0.$$

The three steps above establish that $u_A(p_A,p_B)$ satisfies (3) at all beliefs (p_A,p_B) with $p_A \geq p_B$ and $p_A > {}^c\!/\!_\lambda$.

□

Proof of Proposition 2. If $\gamma \geq \tilde{\gamma}$, then, by definition, given any beliefs (p_A, p_B) with $p_A \geq p_B$ at which no project has been discarded so far, we also have that $f(p_A) \geq p_A$. Hence, the proof verifies that u_D is a solution to the Bellman equation (3) at all such beliefs. First, note that

$$\begin{aligned} V_B[u_D](p_A, p_B) - \frac{c+\gamma}{r} &= \frac{p_B\lambda}{r}\left[1 - u_D(p_A) - (c+\gamma)\right] \\ &\leq u_D(p_A), \end{aligned}$$

where the inequality follows from the fact that $p_B \leq f(p_A)$. Second, it follows from Lemma 1 that

$$\begin{aligned} V_A[u_D](p_A, p_B) - \frac{c+\gamma}{r} &= u_D(p_A) - \frac{\gamma}{r} \\ &< u_D(p_A), \end{aligned}$$

so that (3) is satisfied. □

Proof of Propsition 3. Suppose that $\gamma \in (0, \tilde{\gamma})$, fix beliefs (p_A, p_B) with $\check{p} \geq p_A \geq p_B$, and assume that no project has been discarded so far.

Step 1. It needs to be verified that u_D is a solution to the Bellman equation (3) when $p_B \leq f(p_A)$. First, note that

$$\begin{aligned} V_A[u_D](p_A, p_B) - \frac{c+\gamma}{r} &= u_D(p_A) - \frac{\gamma}{r} \\ &< u_D(p_A). \end{aligned}$$

Second, using $\partial/\partial p_B u_D(p_A) = 0$, we have that

$$\begin{aligned} V_A[u_D](p_A, p_B) - \frac{c+\gamma}{r} - \left[V_B[u_D](p_A, p_B) - \frac{c+\gamma}{r}\right] &= u_D(p_A) - \frac{\gamma}{r} - \frac{1}{r}\left[p_B\lambda[1 - u_D(p_A)] - [c+\gamma]\right] \\ &= u_D(p_A)\left[1 + \frac{p_B\lambda}{r}\right] - \left[\frac{p_B\lambda - c}{r}\right] \\ &> 0, \end{aligned}$$

so that (3) is satisfied.

Step 2. It needs to be verified that u_L is a solution to the Bellman equation (3) when $p_B > f(p_A)$. First, since $u_L(p_A, f(p_A)) = u_D(p_A)$ and $u_L(p_A, p_B)$ is increasing in p_B, it follows that

$$\begin{aligned} V_B[u_L](p_A, p_B) - \frac{c+\gamma}{r} &= u_L(p_A, p_B) \\ &> u_D(p_A). \end{aligned}$$

Second, we have that

$$\begin{aligned}
V_B[u_L](p_A, f(p_A)) - \frac{c+\gamma}{r} - \left[V_A[u_L](p_A, f(p_A)) - \frac{c+\gamma}{r}\right] \\
&= \frac{p_A\lambda + r}{r}\left[u_D(p_A) - \left[\frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}\frac{\partial}{\partial p_A}u_L(p_A, f(p_A))\right]\right] \\
&= \frac{p_A\lambda + r}{r}\left[u_D(p_A) - \left[\frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}u_D'(p_A)\right]\right] \\
&= \frac{\gamma}{r} \\
&> 0. \qquad (16)
\end{aligned}$$

The second equality follows since the value-matching and smooth-pasting conditions (4) and (5) at beliefs $(p_A, f(p_A))$ also implies that a smooth-pasting condition holds with respect to belief p_A, or that

$$\frac{\partial}{\partial p_A}u_L(p_A, f(p_A)) = u_D'(p_A).$$

To see this, differentiate the identity (4) with respect to p_A to obtain

$$u_D'(p_A) = \frac{\partial}{\partial p_A}u_L(p_A, f(p_A)) + f'(p_A)\frac{\partial}{\partial p_B}u_L(p_A, f(p_A)),$$

with the result following from (5). Third, we have that

$$\begin{aligned}
V_B[u_L](p_A, p_A) - \frac{c+\gamma}{r} - \left[V_A[u_L](p_A, p_A) - \frac{c+\gamma}{r}\right] \\
&= \frac{p_A\lambda + r}{r}\left[u_L(p_A, p_A) - \left[\frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}\frac{\partial}{\partial p_A}u_L(p_A, p_A)\right]\right] \\
&= \frac{\lambda p_A(1-p_A)}{r}\left[\frac{\partial}{\partial p_A}u_L(p_A, p_A) - \frac{\partial}{\partial p_B}u_L(p_A, p_A)\right] \\
&\geq 0, \qquad (17)
\end{aligned}$$

where the inequality follows since $p_A \leq \check{p}$. Note that equations (16) and (17) ensure that the Bellman equation (3) is satisfied at beliefs $(p_A, f(p_A))$ and (p_A, p_A). Fourth, to verify that (3) is also satisfied at beliefs (p_A, p_B) with $p_B \in (f(p_A), p_A)$, first note that the function $y(p_B; p_A)$ defined as

$$\begin{aligned}
y(p_B; p_A) &= \frac{r}{p_A\lambda + r}\left[V_B[u_L](p_A, p_B) - \frac{c+\gamma}{r} - \left[V_A[u_L](p_A, p_B) - \frac{c+\gamma}{r}\right]\right] \\
&= C_L(p_A)z(p_B) + p_B\frac{\lambda - [c+\gamma]}{r+\lambda} - (1-p_B)\frac{c+\gamma}{r} \\
&\quad - \left[\frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}C_L'(p_A)z(p_B)\right],
\end{aligned}$$

is convex since $z(\cdot)$ is convex. Furthermore, by the above we have that $y(f(p_A);p_A) > 0$ and $y(p_A;p_A) \geq 0$. Hence, if $y'(p_A;p_A) \leq 0$, it must be that $y(p_B;p_A) \geq 0$, and hence that $V_B[u_L](p_A,p_B) - {}^{(c+\gamma)}/_r - [V_A[u_L](p_A,p_B) - {}^{(c+\gamma)}/_r] \geq 0$, for all $p_B \in (f(p_A), p_A)$. Finally,

$$\begin{aligned} y'(p_A;p_A) &= \frac{\partial}{\partial p_B} u_L(p_A,p_A) + \frac{\lambda p_A(1-p_A)}{p_A\lambda + r} C_L'(p_A) z'(p_A) \\ &= \frac{\partial}{\partial p_B} u_L(p_A,p_A) - C_L'(p_A) z(p_A) \\ &= \frac{\partial}{\partial p_B} u_L(p_A,p_A) - \frac{\partial}{\partial p_A} u_L(p_A,p_A) \\ &\leq 0, \end{aligned}$$

as desired, with the second equality following from the fact that ${}^{z(p)}/_{z'(p)} = -{}^{\lambda p(1-p)}/_{(p\lambda+r)}$.

Step 3. It remains to be shown that u_L satisfying (3) when $p_B > f(p_A)$ and u_D satisfying (3) when $p_B \leq f(p_A)$ is sufficient to ensure that going with the loser and discarding project B, are, respectively, optimal for those regions of beliefs. Note that given any beliefs in $\check{P} = \{(p_A,p_B) : \check{p} \geq p_A \geq p_B\}$, all feasible experimentation policies lead to belief paths that remain in this set, so that the Bellman equation (3) is indeed a necessary and sufficient condition for optimality in the problem with a belief space restricted to $\check{P}$.

□

I make a few remarks before proceeding with the proofs of the results of Section 5.2. For a start, to characterise the mapping g, it will be useful to work with a simple monotone transformation of the difference $V_A[u_L](p_A,p_B) - V_B[u_L](p_A,p_B)$. To this end, define the mapping

$$\begin{aligned} w(p_A,p_B) &= \frac{r}{p_A\lambda + r}\left[V_A[u_L](p_A,p_B) - V_B[u_L](p_A,p_B)\right] \\ &= \frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}\frac{\partial}{\partial p_A} u_L(p_A,p_B) - u_L(p_A,p_B). \end{aligned}$$

It is also useful to state some smooth-pasting properties that the payoff functions u_E and u_W must satisfy. First, note that

$$\begin{aligned} u_W(p_B,p_B) &= \frac{\lambda p_B - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r}\frac{\partial}{\partial p_A} u_W(p_B,p_B) \\ &= u_E(p_B) \\ &= \frac{\lambda p_B - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r}\frac{1}{2} u_E'(p_B), \end{aligned}$$

where the second equality follows from the value-matching condition (7), implies the smooth-pasting condition

$$\frac{\partial}{\partial p_A} u_W(p_B,p_B) = \frac{1}{2} u_E'(p_B), \tag{18}$$

which in turn, since $u_E'(p_B) = {}^{\partial}\!/\!_{\partial p_A} u_W(p_B, p_B) + {}^{\partial}\!/\!_{\partial p_B} u_W(p_B, p_B)$, also implies that

$$\frac{\partial}{\partial p_B} u_W(p_B, p_B) = \frac{1}{2} u_E'(p_B). \tag{19}$$

Second, note that

$$\begin{aligned}
0 &= w(g(p_B), p_B) \\
&= \frac{g(p_B)\lambda - [c+\gamma]}{g(p_B)\lambda + r} - \frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \frac{\partial}{\partial p_A} u_L(g(p_B), p_B) - u_L(g(p_B), p_B) \\
&= \frac{g(p_B)\lambda - [c+\gamma]}{g(p_B)\lambda + r} - \frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \frac{\partial}{\partial p_A} u_L(g(p_B), p_B) - u_W(g(p_B), p_B) \\
&= \frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \left[\frac{\partial}{\partial p_A} u_W(g(p_B), p_B) - \frac{\partial}{\partial p_A} u_L(g(p_B), p_B) \right],
\end{aligned}$$

where the second equality follows from the value-matching condition (10), implies the smooth-pasting condition

$$\frac{\partial}{\partial p_A} u_W(g(p_B), p_B) = \frac{\partial}{\partial p_A} u_L(g(p_B), p_B). \tag{20}$$

The following lemma establishes the key properties of the mappings g and h, as discussed in Section 5.2 and illustrated in Figure 4.

Lemma 2. *If $\underline{p} < \check{p} < \hat{p} < \overline{p}$, then the mappings g and h have the following properties:*

1. *$h(p) > p$ for all $p \in [\check{p}, \hat{p})$ and $h(\hat{p}) = \hat{p}$.*
2. *$g(p_B)$ is decreasing in p_B.*
3. *There exists a belief $\tilde{p} < \check{p}$ such that $g(\tilde{p}) = h(\tilde{p})$. Furthermore, $\tilde{p} > f(g(\tilde{p}))$.*
4. *$h(p_B)$ is increasing in p_B.*

Proof of Lemma 2. For Part 1, I first show that $h(\check{p}) > \check{p}$. To see this, note that while $u_W(\check{p}, \check{p}) = u_L(\check{p}, \check{p})$ by (10), we have that $u_W(p_A, \check{p}) > u_L(p_A, \check{p})$ for all $p_A > \check{p}$ close to $\check{p}$. Indeed, since $C_W(\check{p}) = C_L(\check{p})$, we have that

$$\begin{aligned}
\frac{\partial}{\partial p_A} u_W(\check{p}, \check{p}) &= \frac{\partial}{\partial p_B} u_L(\check{p}, \check{p}) \\
&= \frac{\partial}{\partial p_A} u_L(\check{p}, \check{p}),
\end{aligned}$$

where the second equality follows from the definition of $\check{p}$, but also that

$$\begin{aligned}
\frac{\partial^2}{\partial p_A^2} u_W(\check{p}, \check{p}) &= \frac{\partial^2}{\partial p_B^2} u_L(\check{p}, \check{p}) \\
&> \frac{\partial^2}{\partial p_A^2} u_L(\check{p}, \check{p}),
\end{aligned}$$

as desired. The inequality follows from the fact that $\partial/\partial p_B u_L(p,p) \leq \partial/\partial p_A u_L(p,p)$ for all $p < \check{p}$ close to $\check{p}$ and $\partial/\partial p_B u_L(p,p) > \partial/\partial p_A u_L(p,p)$ for all $p > \check{p}$ close to $\check{p}$. Fourth, a very similar argument establishes that $h(\hat{p}) = \hat{p}$. In this case, we also have that $u_W(\hat{p},\hat{p}) = u_L(\hat{p},\hat{p})$, that $C_W(\hat{p}) = C_L(\hat{p})$ and that $\partial/\partial p_A u_W(\hat{p},\hat{p}) = \partial/\partial p_A u_L(\hat{p},\hat{p})$. However, since $\partial/\partial p_B u_L(p,p) > \partial/\partial p_A u_L(p,p)$ for all $p \leq \hat{p}$ close to $\hat{p}$ and $\partial/\partial p_B u_L(p,p) \leq \partial/\partial p_A u_L(p,p)$ for all $p > \hat{p}$ close to $\hat{p}$, it follows that $\partial^2/\partial p_A^2 u_W(\hat{p},\hat{p}) < \partial^2/\partial p_A^2 u_L(\hat{p},\hat{p})$, so that $u_W(p_A,\hat{p}) < u_L(p_A,\hat{p})$ for all $p_A > \hat{p}$ close to $\hat{p}$. Finally, to show that $h(p) > p$ for all $p \in (\check{p},\hat{p})$, it suffices to show that $u_W(p,p) > u_L(p,p)$ for all such p, given that $p > f(p)$ and hence $u_L(p,p) > u_D(p)$. To see this, fix any $p^* \in [\check{p}, p]$, and, given any belief $p' \in [p^*, p]$, let $u_E^*(p')$ denote the payoff to sharing experimentation from beliefs (p',p') until beliefs (p^*,p^*), after which the experimenter goes with the loser. By the same arguments that accompanied the construction of u_E in Section 5.2, it follows that

$$u_E^*(p') = C_E^* z(p')^2 + p'^2 \frac{\lambda - [c+\gamma]}{r+\lambda} + 2p'(1-p')\frac{\frac{\lambda}{2} - [c+\gamma]}{r+\frac{\lambda}{2}} - (1-p')^2 \frac{c+\gamma}{r},$$

where the constant of integration

$$C_E^* = \frac{u_L(p^*,p^*) - \left[p^{*2}\frac{\lambda-[c+\gamma]}{r+\lambda} + 2p^*(1-p^*)\frac{\frac{\lambda}{2}-[c+\gamma]}{r+\frac{\lambda}{2}} - (1-p^*)^2\frac{c+\gamma}{r}\right]}{z(p^*)^2},$$

is obtained by imposing the value-matching condition

$$u_E^*(p^*) = u_L(p^*,p^*). \tag{21}$$

Note that $\partial/\partial p^* C_E^* \leq 0$ if and only if

$$\begin{aligned}
0 \geq & \left[\frac{\partial}{\partial p_A} u_L(p^*,p^*) + \frac{\partial}{\partial p_B} u_L(p^*,p^*) - \left[u_E^{*\prime}(p^*) - 2C_E^* z'(p^*)z(p^*)\right]\right] z(p^*)^2 \\
& - 2z'(p^*)z(p^*)\left[u_L(p^*,p^*) - \left[u_E^*(p^*) - C_E^* z(p^*)^2\right]\right] \\
= & \left[\frac{\partial}{\partial p_A} u_L(p^*,p^*) + \frac{\partial}{\partial p_B} u_L(p^*,p^*) - u_E^{*\prime}(p^*)\right] z(p^*)^2,
\end{aligned}$$

which holds since $\partial/\partial p_A u_L(p^*,p^*) \leq \partial/\partial p_B u_L(p^*,p^*) = u_E^{*\prime}(p^*)/2$, with the inequality strict if $p^* > \check{p}$. This last inequality follows by definition since $p^* \in [\check{p},\hat{p})$, and the last equality reflects a smooth-pasting property, which can be verified by noting that

$$\begin{aligned}
u_L(p^*,p^*) &= \frac{\lambda p^* - [c+\gamma]}{p^*\lambda + r} - \frac{\lambda p^*(1-p^*)}{p^*\lambda + r}\frac{\partial}{\partial p_B} u_L(p^*,p^*) \\
&= u_E^*(p^*) \\
&= \frac{\lambda p^* - [c+\gamma]}{p^*\lambda + r} - \frac{\lambda p^*(1-p^*)}{p^*\lambda + r}\frac{1}{2} u_E^{*\prime}(p^*),
\end{aligned}$$

where the second equality follows from (21). Finally, since u_E^* is increasing in C_E^* and $p > \check{p}$, it follows that

$$\begin{aligned} u_W(p,p) &= u_E^*(p)\big|_{p^*=\check{p}} \\ &> u_E^*(p)\big|_{p^*=p} \\ &= u_L(p,p), \end{aligned}$$

as desired.

For Part 2, first suppose that $w(g(p_B), p_B) = 0$ in a neighbourhood of p_B. Applying the implicit function theorem and differentiating this identity with respect to p_B yields that

$$\frac{\partial}{\partial p_A} w(g(p_B), p_B) g'(p_B) + \frac{\partial}{\partial p_B} w(g(p_B), p_B) = 0,$$

so that g is differentiable whenever ${\partial}/{\partial p_A} w(g(p_B), p_B) \neq 0$. Second, note that

$$\begin{aligned} \frac{\partial}{\partial p_B} w(g(p_B), p_B) &= -\left[\frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \frac{\partial^2}{\partial p_A \partial p_B} u_L(g(p_B), p_B) + \frac{\partial}{\partial p_B} u_L(g(p_B), p_B)\right] \\ &= -\left[\frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \frac{\partial}{\partial p_A}\left[\frac{p_B\lambda + r}{\lambda p_B(1-p_B)}\left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r} - u_L(g(p_B), p_B)\right]\right]\right. \\ &\qquad \left. + \frac{\partial}{\partial p_B} u_L(g(p_B), p_B)\right] \\ &= \frac{p_B\lambda + r}{\lambda p_B(1-p_B)} \frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \frac{\partial}{\partial p_A} u_L(g(p_B), p_B) - \frac{\partial}{\partial p_B} u_L(g(p_B), p_B) \\ &= \frac{p_B\lambda + r}{\lambda p_B(1-p_B)}\left[\frac{g(p_B)\lambda - [c+\gamma]}{g(p_B)\lambda + r} - \frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r}\right. \\ &\qquad \left. + \frac{\lambda p_B(1-p_B)}{p_B\lambda + r} \frac{\partial}{\partial p_B} u_L(g(p_B), p_B)\right] - \frac{\partial}{\partial p_B} u_L(g(p_B), p_B) \\ &= \frac{p_B\lambda + r}{\lambda p_B(1-p_B)}\left[\frac{g(p_B)\lambda - [c+\gamma]}{g(p_B)\lambda + r} - \frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r}\right] \\ &\geq 0, \end{aligned}$$

with the inequality strict if $g(p_B) > p_B$. The second and fourth equalities follow from (10), and the inequality follows since ${(p\lambda - [c+\gamma])}/{(p\lambda + r)}$ is strictly increasing in p. Third, note that when g is differentiable we have that

$$\begin{aligned} w(g(p_B), p_B) &= \frac{g(p_B)\lambda - [c+\gamma]}{g(p_B) + r} - \frac{\lambda g(p_B)(1-g(p_B))}{g(p_B)\lambda + r} \frac{\partial}{\partial p_A} u_L(g(p_B), p_B) \\ &\qquad - \left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r} \frac{\partial}{\partial p_B} u_L(g(p_B), p_B)\right] \\ &= u_W(g(p_B), p_B) - \left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r} \frac{\partial}{\partial p_B} u_W(g(p_B), p_B)\right], \end{aligned}$$

where the second equality follows by the smooth-pasting properties (20) and

$$\frac{\partial}{\partial p_B}u_W(g(p_B),p_B)=\frac{\partial}{\partial p_B}u_L(g(p_B),p_B). \tag{22}$$

The latter follows by differentiating (10) with respect to p_B to obtain

$$\frac{\partial}{\partial p_B}u_W(g(p_B),p_B)-\frac{\partial}{\partial p_B}u_L(g(p_B),p_B)=g'(p_B)\left[\frac{\partial}{\partial p_A}u_L(g(p_B),p_B)-\frac{\partial}{\partial p_A}u_W(g(p_B),p_B)\right],$$

where (22) follows from (20). Finally, when g is differentiable

$$\begin{aligned}\frac{\partial}{\partial p_A}w(g(p_B),p_B)&=\frac{\partial}{\partial p_A}u_W(g(p_B),p_B)+\frac{\lambda p_B(1-p_B)}{p_B\lambda+r}\frac{\partial^2}{\partial p_B\partial p_A}u_W(g(p_B),p_B)\\&\geq 0,\end{aligned}$$

with the inequality strict if $g(p_B)>p_B$. The inequality follows from an argument almost identical to that establishing the related inequality in the second point of this paragraph. Hence, it follows that $g'(p_B)<0$ whenever $g(p_B)>p_B$ and $\partial/\partial p_A w(g(p_B),p_B)\neq 0$.

For Part 3, first note that $\partial/\partial p_A w(\check{p},\check{p})>0$. To see this suppose, towards a contradiction, that there exists some $p_A>\check{p}$ such that $w(p'_A,\check{p})\leq 0$ for all $p'_A\in[\check{p},p_A]$. For any such p'_A, let $u^*_W(p'_A,\check{p})$ denote the payoff to experimenting with project A until its belief drops to $p^*_A\in[\check{p},p'_A]$, after which the experimenter goes with the loser. By the same arguments that accompanied the construction of u_W in Section 5.2, it follows that

$$u^*_W(p'_A,\check{p})=C^*_W z(p'_A)+p'_A\frac{\lambda-[c+\gamma]}{r+\lambda}-(1-p'_A)\frac{c+\gamma}{r},$$

where the constant of integration

$$C^*_W=\frac{u_L(p^*_A,\check{p})-\left[p^*_A\frac{\lambda-[c+\gamma]}{r+\lambda}-(1-p^*_A)\frac{c+\gamma}{r}\right]}{z(p^*_A)},$$

is obtained by imposing the value-matching condition

$$u^*_W(p^*_A,\check{p})=u_L(p^*_A,\check{p}). \tag{23}$$

Note that $\partial/\partial p^*_A C^*_W\geq 0$ if and only if

$$\begin{aligned}0\leq&\left[\frac{\partial}{\partial p_A}u_L(p^*_A,\check{p})-\left[\frac{\partial}{\partial p_A}u^*_W(p^*_A,\check{p})-C^*_W z'(p^*_A)\right]\right]z(p^*_A)\\&-z'(p^*_A)\left[u_L(p^*_A,\check{p})-[u^*_W(p^*_A,\check{p})-C^*_W z(p^*_A)]\right]\\=&\left[\frac{\partial}{\partial p_A}u_L(p^*_A,\check{p})-\frac{\partial}{\partial p_A}u^*_W(p^*_A,\check{p})\right]z(p^*_A),\end{aligned}$$

which holds since, by assumption,

$$
\begin{aligned}
0 &\geq w(p_A^*, \check{p}) \\
&= \frac{p_A^*\lambda - [c+\gamma]}{p_A^*\lambda + r} - \frac{\lambda p_A^*(1-p_A^*)}{p_A^*\lambda + r}\frac{\partial}{\partial p_A^*} u_L(p_A^*, \check{p}) - u_W(p_A^*, \check{p}) \\
&= \frac{\lambda p_A^*(1-p_A^*)}{p_A^*\lambda + r}\left[\frac{\partial}{\partial p_A^*} u_W(p_A^*, \check{p}) - \frac{\partial}{\partial p_A^*} u_L(p_A^*, \check{p})\right],
\end{aligned}
$$

where the first equality follows from (23). Finally, since u_W^* is increasing in C_W^*, it follows that

$$
\begin{aligned}
u_W(p_A, \check{p}) &= u_W^*(p_A, \check{p})\big|_{p_A^* = \check{p}} \\
&\leq u_W^*(p_A, \check{p})\big|_{p_A^* = p_A} \\
&= u_L(p_A, \check{p}),
\end{aligned}
$$

which contradicts the fact, established in Part 1, that $u_W(p_A, \check{p}) > u_L(p_A, \check{p})$ for all $p_A > \check{p}$ close to $\check{p}$. Second, given that $\partial/\partial p_A w(\check{p}, \check{p}) > 0$, it follows from Part 2 that g is differentiable and strictly decreasing at all beliefs $p \in (\tilde{p}, \check{p}]$, where $\tilde{p} < \check{p}$ is such that $w(p_A, \tilde{p}) \leq 0$ for all $p_A > g(\tilde{p})$ close to $\tilde{p}$. Third, note that $g(\tilde{p}) = h(\tilde{p})$. To see this, suppose, towards a contradiction, that $h(\tilde{p}) > g(\tilde{p})$. This implies that there exists a belief p_A close to $g(\tilde{p})$ such that $u_W(p'_A, g(\tilde{p})) > u_L(p'_A, g(\tilde{p}))$ and $w(p'_A, \tilde{p}) \leq 0$ for all $p'_A \in (g(\tilde{p}), p_A]$. But then, by the argument in the first point of this paragraph, it follows that $u_W(p_A, g(\tilde{p})) \leq u_L(p_A, g(\tilde{p}))$, yielding the desired contradiction. Finally, to see that $\tilde{p} > f(g(\tilde{p}))$, note that, given any beliefs $(p_A, f(p_A))$,

$$
\begin{aligned}
w(p_A, f(p_A)) &= \frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}\frac{\partial}{\partial p_A} u_L(p_A, f(p_A)) - u_L(p_A, f(p_A)) \\
&= \frac{p_A\lambda - c}{p_A\lambda + r} - \frac{\lambda p_A(1-p_A)}{p_A\lambda + r} u'_D(p_A) - u_D(p_A) - \frac{\gamma}{p_A\lambda + r} \\
&= -\frac{\gamma}{p_A\lambda + r} \\
&< 0,
\end{aligned}
$$

where the second equality follows from (4) and (5), and the final equality follows since $u_D(p_A) = (p_A\lambda - c)/(p_A\lambda + r) - \lambda p_A(1-p_A)/(p_A\lambda + r) u'_D(p_A)$. Hence, it cannot be that $p_A = f(g(p_B))$.

For Part 4, first suppose that $u_W(h(p_B), p_B) = u_D(h(p_B))$ in a neighbourhood of p_B. Differentiating this identity yields that

$$
h'(p_B)\left[u'_D(h(p_B)) - \frac{\partial}{\partial p_A} u_W(h(p_B), p_B)\right] = \frac{\partial}{\partial p_B} u_W(h(p_B), p_B),
$$

where $\partial/\partial p_B u_W(h(p_B), p_B) > 0$. By rearranging the identity $u_W(h(p_B), p_B) = u_D(h(p_B))$, we obtain that

$$
C_W(p_B) - C_D = \frac{\gamma}{z(h(p_B))}\frac{r + \lambda(1 - h(p_B))}{r(r+\lambda)},
$$

and so

$$u'_D(h(p_B)) - \frac{\partial}{\partial p_A} u_W(h(p_B), p_B) = z'(h(p_B))\left[C_D - C_W(p_B)\right] - \frac{\lambda\gamma}{r(r+\lambda)}$$
$$= \frac{\gamma}{r(r+\lambda)}\left[\frac{[h(p_B)\lambda + r][r + \lambda(1 - h(p_B))]}{\lambda h(p_B)(1 - h(p_B))} - \lambda\right]$$
$$> 0,$$

so that h is differentiable, and $h'(p_B) > 0$, as desired. The second equality follows since ${}^{z(p)}/_{z'(p)} = -{}^{\lambda p(1-p)}/_{(p\lambda+r)}$. Second, suppose that $u_W(h(p_B), p_B) = u_L(h(p_B), p_B)$ in a neighbourhood of p_B. Differentiating this identity yields that

$$h'(p_B)\left[\frac{\partial}{\partial p_A} u_L(h(p_B), p_B) - \frac{\partial}{\partial p_A} u_W(h(p_B), p_B)\right] = \frac{\partial}{\partial p_B} u_W(h(p_B), p_B) - \frac{\partial}{\partial p_B} u_L(h(p_B), p_B).$$

Third, the inequality (25), which I establish in Step 2 of the proof of Proposition 4, shows that ${}^{\partial}/_{\partial p_A} u_L(h(p_B), p_B) \geq {}^{\partial}/_{\partial p_A} u_W(h(p_B), p_B)$. Fourth, I also establish in Step 3 of the proof of Proposition 4 that $V_A[u_W](h(p_B), p_B) - {}^{(c+\gamma)}/_r \geq V_B[u_W](h(p_B), p_B) - {}^{(c+\gamma)}/_r$, which implies that ${}^{\partial}/_{\partial p_B} u_W(h(p_B), p_B) \geq {}^{\partial}/_{\partial p_B} u_L(h(p_B), p_B)$, with the inequality strict if $h(p_B) > p_B$. Fifth, remember that by Part 1, we have that $h(p_B) > p_B$ for all $p_B \in [\check{p}, \hat{p})$, while by Parts 2 and 3 we have that $h(p_B) \geq g(p_B) > p_B$ for all $p_B \in [\tilde{p}, \check{p})$. Hence, restricting attention to $p_B \in [\tilde{p}, \hat{p})$, if ${}^{\partial}/_{\partial p_A} u_W(h(p_B), p_B) > {}^{\partial}/_{\partial p_A} u_L(h(p_B), p_B)$, then $h(p_B)$ is differentiable and $h'(p_B) > 0$, as desired. Suppose instead that ${}^{\partial}/_{\partial p_A} u_W(h(p_B), p_B) = {}^{\partial}/_{\partial p_A} u_L(h(p_B), p_B)$. In that case, we have both that ${}^{\partial}/_{\partial p_B} u_W(g(p_B), p_B) > {}^{\partial}/_{\partial p_B} u_L(g(p_B), p_B)$, as well as $u_W(p_A, p_B) > u_L(p_A, p_B)$ for all $p_A < h(p_B)$. Since $h(p_B)$ is continuous, it must be that $h(p'_B) > h(p_B)$ for all $p'_B > p_B$ close to p_B, as desired.

□

Proof of Proposition 4. Suppose that $\gamma \in (0, \tilde{\gamma})$, fix beliefs (p_A, p_B) with $p_A \geq p_B$ and $p_A \in (\check{p}, \hat{p}]$, and assume that no project has been discarded so far.

Step 1. It needs to be verified that for all $(p_A, p_B) \notin \mathcal{P}_W$ with $p_B \leq f(p_A)$, u_D is a solution to the Bellman equation (3). This is the same as Step 1 in the proof of Proposition 3.

Step 2. It needs to be verified that for all $(p_A, p_B) \notin \mathcal{P}_W$ with $p_B > f(p_A)$, u_L is a solution to the Bellman equation (3). First, as noted in Step 2 of the proof of Proposition 3, $V_B[u_L](p_A, p_B) - {}^{(c+\gamma)}/_r > u_D(p_A)$ follows from $p_B > f(p_A)$. Second, as noted in Step 2 of the proof of Proposition 3, $V_B[u_L](p_A, f(p_A)) - {}^{(c+\gamma)}/_r > V_A[u_L](p_A, f(p_A)) - {}^{(c+\gamma)}/_r$. Third, there are three cases, either (i) $p_A \in [\hat{p}, \mathring{p})$, in which case $(p_A, p_A) \notin \mathcal{P}_W$ and $p_A > f(p_A)$, (ii) there exist (p_A, p'_B) with $p'_B > p_B$ and $p_A = g(p'_B)$, or (iii) there exist (p_A, p'_B) with $p'_B > p_B$ and $p_A = h(p'_B)$. As noted in Step 2 of the proof of Proposition 3, for case (i) $V_B[u_L](p_A, p_A) - {}^{(c+\gamma)}/_r \geq V_A[u_L](p_A, p_A) - {}^{(c+\gamma)}/_r$ follows from the fact that $p_A \in [\hat{p}, \mathring{p})$ and hence that ${}^{\partial}/_{\partial p_A} u_L(p_A, p_A) \geq {}^{\partial}/_{\partial p_B} u_L(p_A, p_A)$. For case

(ii), note that

$$
\begin{aligned}
&V_B[u_L](g(p'_B),p'_B) - \frac{c+\gamma}{r} - \left[V_A[u_L](g(p'_B),p'_B) - \frac{c+\gamma}{r}\right] \\
&\quad = \frac{g(p'_B)\lambda + r}{r}\Bigg[u_L(g(p'_B),p'_B) - \Bigg[\frac{g(p'_B)\lambda - [c+\gamma]}{g(p'_B)\lambda + r} \\
&\qquad\qquad - \frac{\lambda g(p'_B)(1-g(p'_B))}{g(p'_B)\lambda + r}\frac{\partial}{\partial p_A}u_L(g(p'_B),p'_B)\Bigg]\Bigg] \\
&\quad = 0,
\end{aligned} \tag{24}
$$

which follows from (10) and (20). For case (iii), note that

$$
\begin{aligned}
&V_B[u_L](h(p'_B),p'_B) - \frac{c+\gamma}{r} - \left[V_A[u_L](h(p'_B),p'_B) - \frac{c+\gamma}{r}\right] \\
&\quad = \frac{h(p'_B)\lambda + r}{r}\Bigg[u_W(h(p'_B),p'_B) - \Bigg[\frac{h(p'_B)\lambda - [c+\gamma]}{h(p'_B)\lambda + r} \\
&\qquad\qquad - \frac{\lambda h(p'_B)(1-h(p'_B))}{h(p'_B)\lambda + r}\frac{\partial}{\partial p_A}u_L(h(p'_B),p'_B)\Bigg]\Bigg] \\
&\quad = \frac{\lambda h(p'_B)(1-h(p'_B))}{r}\left[\frac{\partial}{\partial p_A}u_L(h(p'_B),p'_B) - \frac{\partial}{\partial p_A}u_W(h(p'_B),p'_B)\right] \\
&\quad \geq 0.
\end{aligned} \tag{25}
$$

The inequality follows from the fact, by the definition of $h(p'_B)$, $u_L(p'_A,p'_B) < u_W(p'_A,p'_B)$ for all $p'_A < h(p'_B)$ close to $h(p'_B)$. Fourth, as argued in Step 2 of the proof of Proposition 3, to show that $V_B[u_L](p_A,p''_B) - {}^{(c+\gamma)}/_r \geq V_B[u_L](p_A,p''_B) - {}^{(c+\gamma)}/_r$ for all $p''_B \in [p_B,p'_B)$, it is sufficient to show that, in cases (ii) and (iii), $y'(p'_B;p_A) \leq 0$ (as the corresponding claim for case (i) is dealt with there). Note that

$$
\begin{aligned}
y'(p'_B;p_A) &= \frac{\partial}{\partial p_B}u_L(p_A,p'_B) + \frac{\lambda p_A(1-p_A)}{p_A\lambda + r}\frac{\partial^2}{\partial p_A \partial p_B}u_L(p_A,p'_B) \\
&= -\frac{p'_B\lambda + r}{\lambda p'_B(1-p'_B)}\left[\frac{p_A\lambda - [c+\gamma]}{p_A\lambda + r} - \frac{p'_B\lambda - [c+\gamma]}{p'_B\lambda + r}\right] \\
&\leq 0,
\end{aligned}
$$

which follows from the same manipulations as in the proof of Part 2 of Lemma 2, since in case (ii) we have that $p_A = g(p'_B)$ and $u_W(g(p'_B),p'_B) = u_L(g(p'_B),p'_B)$, and in case (iii) we have that $p_A = h(p'_B)$ and $u_W(h(p'_B),p'_B) = u_L(h(p'_B),p'_B)$. Hence, (3) is satisfied.

Step 3. It needs to be verified that for all $(p_A,p_B) \in \mathcal{P}_W$, u_W is a solution to the Bellman equation (3). First, since $u_W(p_A,p_B) \geq \max\{u_L(p_A,p_B),u_D(p_A)\}$ by definition, it follows that

$$
\begin{aligned}
V_A(p_A,p_B) - \frac{c+\gamma}{r} &= u_W(p_A,p_B) \\
&\geq u_D(p_A).
\end{aligned}
$$

Second, there are two cases, either (i) $(p_B, p_B) \in \mathcal{P}_W$, in which case $p_B \in [\check{p}, \hat{p})$, or (ii) there exist beliefs $(g(p_B), p_B)$, in which case $p_B < \check{p}$. In case (i), we have that

$$\begin{aligned}
V_A[u_W](p_B, p_B) - \frac{c+\gamma}{r} - &\left[V_B[u_W](p_B, p_B) - \frac{c+\gamma}{r}\right] \\
&= -\frac{\lambda p_B(1-p_B)}{r}\left[\frac{\partial}{\partial p_A} u_W(p_B, p_B) - \frac{\partial}{\partial p_B} u_W(p_B, p_B)\right] \\
&= 0,
\end{aligned} \tag{26}$$

where the second equality follows from (18) and (19). In case (ii), we have that

$$\begin{aligned}
V_A[u_W](g(p_B), p_B) - \frac{c+\gamma}{r} - &\left[V_B[u_W](g(p_B), p_B) - \frac{c+\gamma}{r}\right] \\
&= \frac{\lambda p_B + r}{r}\left[u_W(g(p_B), p_B) - \left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r}\right.\right. \\
&\qquad \left.\left. - \frac{\lambda p_B(1-p_B)}{p_B\lambda + r}\frac{\partial}{\partial p_B} u_W(g(p_B), p_B)\right]\right] \\
&= \frac{\lambda p_B + r}{r}\left[u_W(g(p_B), p_B) - u_L(g(p_B), p_B)\right] \\
&= 0,
\end{aligned}$$

where the second equality follows the smooth-pasting property (22) established in the proof of Part 2 of Lemma 2. Third, in case (i), we have that

$$\begin{aligned}
&\frac{r}{\lambda p_B + r} \cdot \frac{\partial}{\partial p_A}\left[V_A[u_W](p_B, p_B) - \frac{c+\gamma}{r} - \left[V_B[u_W](p_B, p_B) - \frac{c+\gamma}{r}\right]\right] \\
&\quad = \frac{\partial}{\partial p_A}\left[u_E(p_B) - \left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{\lambda p_B + r}\frac{1}{2}u_E'(p_B)\right]\right] \\
&\quad = \frac{\partial}{\partial p_A}\left[u_E(p_B) - u_E(p_B)\right] \\
&\quad = 0,
\end{aligned}$$

where the first equality follows from the value-matching and smooth-pasting conditions (9), (18) and (19), while in case (ii), we have that

$$\begin{aligned}
&\frac{r}{\lambda p_B + r} \cdot \frac{\partial}{\partial p_A}\left[V_A[u_W](g(p_B), p_B) - \frac{c+\gamma}{r} - \left[V_B[u_W](g(p_B), p_B) - \frac{c+\gamma}{r}\right]\right] \\
&\quad = \frac{\partial}{\partial p_A}\left[u_W(g(p_B), p_B) - \left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{\lambda p_B + r}\frac{\partial}{\partial p_B} u_W(g(p_B), p_B)\right]\right] \\
&\quad \geq 0,
\end{aligned}$$

with the inequality strict if $p_A > p_B$, where the inequality follows from the argument that $\partial/\partial p_A w(g(p_B), p_B) \geq 0$ in the proof of Part 2 of Lemma 2. Fourth, note that the mapping

$$\begin{aligned} p_A \mapsto V_A[u_W](p_A, p_B) - \frac{c+\gamma}{r} - \left[V_B[u_W](p_A, p_B) - \frac{c+\gamma}{r}\right] \\ = \frac{\lambda p_A + r}{r}\left[C_W z(p_A) + p_A \frac{\lambda - [c+\gamma]}{r+\lambda} - (1-p_A)\frac{c+\gamma}{r}\right. \\ \left. - \left[\frac{p_B\lambda - [c+\gamma]}{p_B\lambda + r} - \frac{\lambda p_B(1-p_B)}{\lambda p_B + r} C'_W(p_B) z(p_A)\right]\right] \end{aligned}$$

is convex since z is convex. Hence, together with the second and third points, this implies that $V_A[u_W](p_A, p_B) - (c+\gamma)/r \geq V_B[u_W](p_A, p_B) - (c+\gamma)/r$ for all $p_A \geq p_B$ in case (i) and for all $p_A \geq g(p_B)$ in case (ii), with the inequality strict if $p_A > p_B$, so that (3) is satisfied.

Step 4. It remains to be shown that Steps 1-3 above, along with the results of Proposition 3, are sufficient to ensure that the payoff function constructed on the belief space $\mathring{P} = \{(p_A, p_B) : \mathring{p} > p_A \geq p_B\}$ is indeed the value function. But this is the same argument as in the proof of Step 3 of Proposition 3: given any beliefs in $\mathring{P}$, all feasible experimentation policies lead to belief paths that remain in this set, so that the Bellman equation (3) is indeed a necessary and sufficient condition for optimality in the problem with a belief space restricted to $\mathring{P}$.

□