WORKING PAPER SERIES

Managing Multiple Research Projects

Alejandro Francetich

Working Paper n. 516

This Version: October 7, 2014

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Managing Multiple Research Projects*

(JOB-MARKET PAPER)

Alejandro Francetich †‡

Department of Decision Sciences and IGIER, Bocconi University, Italy

October 7, 2014

Abstract

A decision maker can experiment on up to two alternatives *simultaneously* over time. One and only one of these alternatives can produce successes, according to a Poisson process with known arrival rate; but there is uncertainty as to which alternative is the profitable one. The decision maker only observes the outcomes of the alternatives chosen, and choosing each alternative entails a cost. Simultaneous experimentation involves higher costs but can produce more data. At the same time, since the alternatives are negatively correlated, the outcomes of either one are informative about the other. If the costs are high and she is sufficiently impatient, the decision maker never experiments on both alternatives at once. Otherwise, if she starts with a single alternative that produces no successes, she becomes gradually pessimistic and eventually takes on the other alternative *while keeping the first one* — despite the higher costs and the negative correlation.

Keywords: Experimentation, two-armed bandits, multi-choice bandits, negatively correlated arms, Poisson process

JEL Classification Numbers: D83, D90

---

*A previous version of this paper has circulated under the title "Experimentation With Menus."

†Via Röntgen 1, 20136 Milan, Italy. Email address: alejandro.francetich@unibocconi.it

‡This paper borrows inspiration from Chapter 3 of my dissertation, submitted to the Stanford Graduate School of Business. I am deeply indebted to David Kreps for continuously enlightening me, and to Alejandro Manelli, Pierpaolo Battigali, and Massimo Marinacci for their support and guidance. I am grateful to Sven Rady for his encouragement and his insightful and helpful comments. The paper has also benefited from comments by Amanda Friedenberg, Marco Ottaviani, and attendants of the 2014 Canadian Economic Theory Conference and the 25th International Game Theory Conference at Stony Brook. I gratefully acknowledge financial support from ERC advanced grant 324219. Any remaining errors and omissions are all mine.

1 Introduction

A researcher has a conjecture for a new theoretical result. If the conjecture turns out to be true, and if she can provide a proof, she can write a paper with her result. If the conjecture turns out to be incorrect, and she produces a counterexample, she can write a different paper with the "modified conjecture," or perhaps a (shorter) paper with the counterexample. She can divide her time between trying to come up with a proof and trying to come up with a counterexample; but she can also use part of her research funds to hire a research assistant to work on a counterexample *while* she works on the proof, or viceversa. This way, a paper may be produced faster; but research funds have to be spent, and in the end one and only one of the two endeavours can succeed.

Alternatively, a lab is conducting clinical research on two different new treatments for a disease. One of the treatments is based on the hypothesis that the cause of the disease is a virus, while the other, bacteria. The lab director can have her staff experiment on either treatment; or she can hire additional researchers, and have two teams work side by side on the separate treatments. This way, the successful treatment may be identified faster; but the additional researchers must be compensated for their work, and ultimately only one team can be successful.

Similarly, a professional-services firm, such as a consultancy or a legal partnership, employs experts to serve their clients. The cases they handle may involve expertise in, say, regulation or finance for a consultancy, and taxes or litigation for a legal partnership. The manager of the firm is uncertain about the clientele. She can hire an expert on one of these areas, or hire multiple experts, one in each field. At the cost of a higher payroll, the firm can be better equipped to handle cases and get more data on their clients' profiles. But experts that remain idle during any period must nonetheless be paid a wage.

This paper studies such experimentation problems, where a decision maker faces two alternatives (research projects to work on, treatments to try, experts to hire) and two possible, ex-ante unobserved states of nature (truth value of a research conjecture, of a medical hypothesis, client profiles). Successes arrive over time according to a known Poisson process, but correspond to one and only one of these alternatives according to the unobserved state of nature. The alternatives are thus negatively correlated: A success from one is conclusive evidence that the other alternative cannot produce value. Over any time interval, the decision maker can choose to not experiment, to experiment on a single alternative, or to experiment on both alternatives *at once*. She bears a (constant) flow cost for each alternative chosen, and she only observes successes *for those alternatives selected*. Therefore, by experimenting on a single alternative, she cannot distinguish between the arrival of a success for the other alternative and failure of arrival altogether; she must experiment on both *simultaneously* to tell these two events apart. Nonetheless, the decision maker can make inferences about any alternative from the outcomes of the other alternative, due to their correlation.

Imagine that the decision maker can only experiment on at most one project at a time. If the cost of experimentation is too high, and if she is sufficiently impatient (alternatively, if the arrival rates of successes are too low), the decision maker starts on the project that she deems most promising. As long as it fails, she gradually loses confidence in the chosen alternative, in favor of the neglected one: Bad news about a project is good news about the other project. Eventually, however, the decision maker gives up altogether: Before her posterior can reach a point where it would be optimal to switch projects, it reaches a point of uncertainty at which no further experimentation is worthwhile. Intuitively, the continued failure of her initially favored project has left her unsure — the odds have not been turned far enough —, and further experimentation is simply not worth her while.

The basic results when simultaneous experimentation is allowed are summarized as follows:

(a) If the costs of the projects are low and/or the arrival rates are sufficiently high relative to the discount rate, the decision maker begins with both projects at once when her prior is diffuse. If, instead, her prior assesses that one project is sufficiently likely to be the fruitful one, she begins with that project alone. If enough time passes and she meets with no success, she takes on *both projects simultaneously* once she becomes sufficiently unsure about the state; she never abandons one project for the other.

(b) But if the costs of the projects are high, and if the discount rate is high (or the arrival rate is low), she either does no research at all — if her prior is sufficiently diffuse — or she works on one project only — if the prior that said project is the fruitful project is high enough —, abandoning research if, after a while, she does not meet with success.

(c) Imagine the researcher, if she conducts both projects at once and success is achieved, cannot tell which project was responsible for the success. The manager of the consulting firm may observe whether a team of experts collectively meets their clients' needs, but not exactly how much each of the experts contributes individually. For low costs and high discount rate, the decision maker starts with both projects if her prior is diffuse. Now, however, if she starts on a singleton, she sticks to the singleton for longer. If the costs are too high relative to the arrival rate, or if the discount rate is sufficiently low, she only experiments on singletons: Information can only come from singletons, which are cheaper than simultaneous experimentation.

(d) Of the two basic alternatives, one must eventually succeed; in this sense, they are "collectively safe." Sometimes, a researcher may also have other, separate projects to work on, projects that may fail. The decision maker postpones starting on these collectively-safe projects in favor of a riskier one if she is sufficiently optimistic about this third project, less optimism required the more uncertainty there is about the two original projects.

The decision maker experiments simultaneously on both projects when she is sufficiently uncertain about the state. In this sense, there is "more experimentation" for mid-range beliefs, when information is the most valuable. This stands in (apparent) contrast with Moscarini and Smith (2001), who find that experimentation "accelerates" when the decision maker is close to being confident enough to make a choice.

Moscarini and Smith (2001) assume that the cost of "experimentation" — in their model, buying signals and delaying final, irreversible choices — is strictly increasing and strictly convex.

Moreover, observations and posteriors obey a diffusion process, so they always change gradually over time. Thus, experimentation is more costly when it takes longer for the posterior to reach decision thresholds. The same is true in the present paper: Experimentation is more costly when it has the smallest impact on beliefs. However, the flow cost of choosing each alternative is constant, and observing an arrival from a Poisson process produces jumps in the posterior rather than gradual changes; thus, experimentation here has the smallest impact on beliefs when the posterior is already close to the extremes.

While Moscarini and Smith (2001) represent experimentation as a type of Wald sequential hypothesis-testing problem, I follow the more traditional literature and represent experimentation as a bandit problem — here, a Poisson bandit with correlated arms.[1] Not being restricted to choosing at most a single arm at a time, it is in fact a *multi-choice* bandit problem. A special class of such problems is studied in Bergemann and Välimäki (2001). A decision maker faces countably many arms, and can choose up to some fixed number of them at a time, at no additional cost. A generalization of the Gittins index[2] applies if the arms are independent, ex-ante identical, and there are (countably) infinitely many of them; however, Bergemann and Välimäki (2001) show by example that this solution fails if there are only finitely many arms.

In Francetich and Kreps (2014), we study the following variation of the present multi-choice bandit problem. A finite set X of alternatives, or "tools," is given. Each time period $t = 0, 1, \ldots$, a decision maker chooses a "toolkit" $K_t \subseteq X$ to carry for that period. Each tool $x \in X$ has a "rental" cost $c_x > 0$ and value on date t given by $v_t(x)$, where the process $\{v_t \in R_+^X\}_{t \in \mathbb{N}}$ is independent and identically distributed according to some unknown distribution. On each date t, the decision maker only observes the values $v_t(x)$ for those x that are in the toolkit she has selected, K_t. As a bandit problem with non-independent arms, we cannot enlist the Gittins index. In fact, at this level of generality, the best we can hope for — aside from asymptotic or long-run results —, and the subject of Francetich and Kreps (2014), is to investigate the performance of various decision heuristics. We borrow from the machine-learning literature in computer science and operations research, which is concerned with developing algorithms that "perform well" in bandit problems.[3] But one can imagine special and restricted formulations of this problem that are amenable to analytical solution, and the present paper provides one such formulation; this allows us to build up our intuition regarding solutions to the more general problem.

While the spirit of the problem studied in this paper is closely related to Francetich and Kreps (2014), the formal techniques employed borrow heavily from Keller and Rady (2010) and Klein and Rady (2011), who study *strategic* experimentation with Poisson and exponential bandits, respectively. In Keller and Rady (2010), each player has an identical copy of a bandit with one

[1]In terms of behavior, the Wald approach decouples payoffs and learning, but makes decisions irreversible. In the bandit approach, choices yield both information and payoffs (assessed on the basis of said information), and such choices are typically reversible. I consider some very special forms of irreversibility in section 5.

[2]See, for instance, Gittins and Jones (1974); Whittle (1980); Weber (1992).

[3]References to this literature are provided in Francetich and Kreps (2014). Part of this literature provides algorithms even for problems where the arms of the bandit are statistically independent under the prior, so that the Gittins index can be applied; this is because the computation of the index is typically complex as a practical matter.

risky arm and one safe arm; in Klein and Rady (2011), the risky arms of each player's bandit are negatively correlated *across players*. These players can choose only one arm at a time, but they can learn from each other. Like the decision maker in the present paper, they have more than one source of information. However, to them, this "extra" information is public and free; our decision maker can only exploit her additional source of information — choosing more than one arm at a time— by means of payoff-relevant actions, so she faces a different trade-off.

Nonetheless, the baseline problem of our decision maker can be mapped to the benchmark problem of the social planner in Klein and Rady (2011). From the point of view of this planner, the two agents are our two projects; the opportunity cost of neglecting the safe arm is the cost of the projects; and, since actions and outcomes are public, assigning both agents to the risky arm corresponds to simultaneous experimentation. Thus, the results in section 4 have exact parallels to results in Klein and Rady (2011), and the analysis in section 5 constitutes an extension of their efficiency benchmark.

The rest of the paper is organized as follows. Section 2 describes the basics of the formal framework. Section 3 analyzes the single-choice benchmark, while section 4 studies the baseline multi-choice problem. Section 5 analyzes several extensions and variations of the baseline model: the case of asymmetric costs (section 5.1); the case when choices must be nested (section 5.2); the case of "imperfect monitoring," namely, when successes from simultaneous experimentation cannot be attributed to individual projects (section 5.3); and the case where the decision maker has a third, separate project to try (section 5.4). Finally, section 6 concludes. Proofs are relegated to the appendix, although some derivations that (hopefully) contribute to the exposition are kept in the main body of the text.

2 The Model

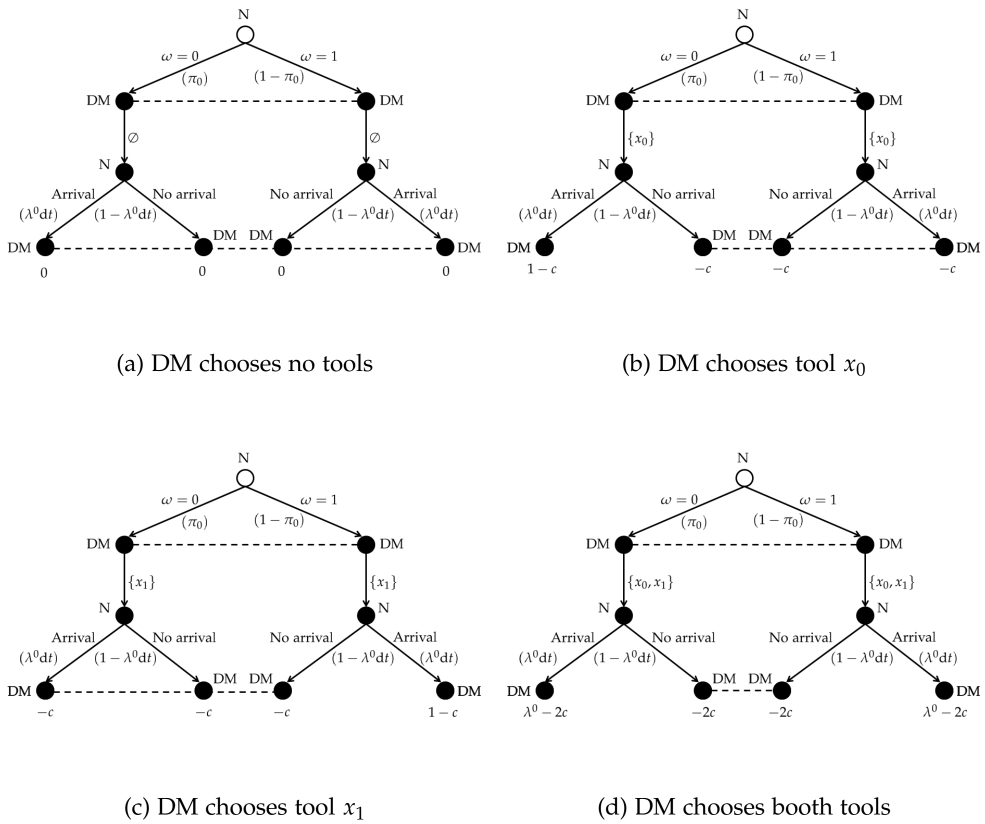
2.1 Framework

There is a set of alternatives $X = \{x_0, x_1\}$, which represents projects, experts, or "tools" a decision maker (henceforth, DM) can experiment on, or employ. The DM allocates her time between the different subsets of X, representing research agendas, teams of experts, or "toolkits." The set of allocations of a unit of time between the subsets $\{x_0\}, \{x_1\}$, and $\{x_0, x_1\}$ is $A := \{\alpha \in [0,1]^3 : \alpha_0 + \alpha_1 + \alpha_2 \leq 1\}$, where α_0 denotes the fraction of a unit of time spent on $\{x_0\}$, α_1 is the fraction of time spent on $\{x_1\}$, and α_2 is the time spent on simultaneous experimentation, while $1 - \alpha_0 - \alpha_1 - \alpha_2$ is the fraction of time spent "on the empty set," namely, doing no research.

Tools must be "rented" to be employed; $c > 0$ is the per-tool rental rate (research costs, wages, fees). If successful, a tool employed yields a gross reward of 1. Successes arrive over time for tools x_0 and x_1 according to Poisson processes with arrival rates $\lambda^0(1-\omega)$ and $\lambda^0\omega$, respectively, where $\lambda^0 > c$ is the *known* arrival rate and $\omega \in \{0, 1\}$ is the ex-ante *unobserved* state of nature. In words, it is known that one and only one of these tools is profitable — and exactly how profitable it is —, but there is uncertainty as to which one is the profitable one.

If $\pi \in [0,1]$ represents the assessment of the DM that $\omega = 0$, rewards are 0 from choosing the empty set; $\lambda^0\pi - c$ from choosing $\{x_0\}$; $\lambda^0(1-\pi) - c$ from choosing $\{x_1\}$; and $\lambda^0 - 2c$ from choosing X. Future payoffs are discounted at the subjective rate $\rho > 0$. Choices also affect the amount of data that the DM collects over any time interval. She only observes the arrival of opportunities for the tools selected; by choosing a single tool, she cannot distinguish between the event of an arrival for the tool she ignored and the event of failure of arrival altogether. Figure 1 summarizes ex-post payoffs and data collected under each of the possible choices.

A more flexible specification would allow for $\omega \in (0,1)$, so that successes can arrive for both tools. Under this alternative specification, successes are "allocated" to tool x_0 or x_1 with probabilities $1-\omega$ and ω, respectively, independently of past arrivals and allocations; this yields a partitioning of the Poisson process of success arrivals. But this additional flexibility comes at the cost of slowing down the learning process, without providing significant new insights. An arrival for a tool ceases to be conclusive evidence that the tool is the superior one. Instead, we would assess a tool to be superior by observing a sufficiently larger frequency of arrivals for it relative to the other tool; a single observation of success no longer suffices.



(a) DM chooses no tools

(b) DM chooses tool x_0

(c) DM chooses tool x_1

(d) DM chooses booth tools

Figure 1: DM's observations and payoffs under each of her possible choices.

2.2 Bayesian Updating and the Bellman Equation

The prior of the DM that $\omega = 0$ is denoted by $\pi_0 \in (0, 1)$; her corresponding posterior at the beginning of period t is denoted by π_t. Given beliefs represented by $\pi \in [0, 1]$, expected immediate rewards from allocating a unit of time according to $\alpha \in A$ are:

$$I(\alpha, \pi, \mathrm{d}t) := \alpha_0 \left(\lambda^0 \pi - c\right) \mathrm{d}t + \alpha_1 \left(\lambda^0 (1 - \pi) - c\right) \mathrm{d}t + \alpha_2 (\lambda^0 - 2c) \mathrm{d}t.$$

The event of arrival makes the posterior jump to 1, if the arrival is from x_0, or to 0, if the arrival is from x_1. By spending time on both tools, either nothing new is learned, or the model uncertainty is resolved immediately. If no arrival results from spending a fraction α_t of time on tool x_0 over the time interval $[t, t + \Delta t)$, the posterior is:

$$\pi_{t+\Delta t} = \frac{\pi_t e^{-\alpha_t \lambda^0 \Delta t}}{\pi_t e^{-\alpha_t \lambda^0 \Delta t} + 1 - \pi_t}.$$

As Δt shrinks, we obtain:

$$\dot{\pi}_t = -\alpha_t \lambda^0 \pi_t (1 - \pi_t).$$

If no arrival results from spending a fraction β_t of time on x_1, we get $\dot{\pi}_t = \beta_t \lambda^0 \pi_t (1 - \pi_t)$.

The problem is stationary, and the state of the problem is the belief of the DM, $\pi \in [0, 1]$.[4] Let $w : [0, 1] \to \mathbb{R}$ denote the (optimal, average) value function. The expected continuation value given $\alpha \in A$ is:

$$\begin{aligned} C(\alpha, \pi, \mathrm{d}t) := {} & (\alpha_0 + \alpha_2)\, \lambda^0 \pi \,\mathrm{d}t \left[w(1) - w(\pi)\right] + (\alpha_1 + \alpha_2)\, \lambda^0 (1 - \pi) \mathrm{d}t \left[w(0) - w(\pi)\right] \\ & + w(\pi) + \left[1 - (\alpha_0 + \alpha_2) \lambda^0 \pi \,\mathrm{d}t - (\alpha_1 + \alpha_2) \lambda^0 (1 - \pi) \mathrm{d}t\right] (\alpha_1 - \alpha_0) \lambda^0 \pi (1 - \pi) w'(\pi) \mathrm{d}t, \end{aligned}$$

where $w(0) = w(1) = \lambda^0 - c$. The Bellman equation of the problem is:

$$w(\pi) = \max_{\alpha \in A} \left\{ \rho I(\alpha, \pi, \mathrm{d}t) + e^{-\rho \mathrm{d}t} C(\alpha, \pi, \mathrm{d}t) \right\}.$$

By invoking the approximations $e^{-\rho \mathrm{d}t} \approx 1 - \rho \mathrm{d}t$ and $(\mathrm{d}t)^n \approx 0$ for all naturals $n \geq 2$, and rearranging terms, we can rewrite the Bellman equation as:

$$\begin{aligned} w(\pi) = \max_{\alpha \in A} \Bigg\{ & \alpha_0 \left[\lambda^0 \pi - c + \frac{\lambda^0 \pi \left(w(1) - w(\pi)\right) - \lambda^0 \pi (1 - \pi) w'(\pi)}{\rho} \right] \\ & + \alpha_1 \left[\lambda^0 (1 - \pi) - c + \frac{\lambda^0 (1 - \pi) \left(w(0) - w(\pi)\right) + \lambda^0 \pi (1 - \pi) w'(\pi)}{\rho} \right] \\ & + \alpha_2 \left[\lambda^0 - 2c + \frac{\lambda^0 \left(\lambda^0 - c - w(\pi)\right)}{\rho} \right] \Bigg\}. \end{aligned}$$

[4]The state space will still be $[0, 1]$ if ω can take two interior values, $0 < \underline{\omega} < \overline{\omega} < 1$. However, accommodating a richer set of states of nature would require specifying a multi-dimensional state space.

Since the expression in braces in the Bellman equation is linear in α, optimal strategies will involve spending the full unit of time on the most promising toolkit, except perhaps at indifference points. Because of the stationarity of the problem, in looking for optimal strategies, we may restrict attention to *stationary strategies*, namely, strategies that recommend toolkits as a function of beliefs. The next theorem establishes that we may further restrict attention to *cutoff strategies*; namely, to stationary strategies with the following properties:

- If the strategy recommends the toolkit $\{x_1\}$ for some $\pi \in [0,1]$, then it also recommends $\{x_1\}$ for any $\pi' \in [0,\pi)$;
- If the strategy recommends choosing $\{x_0\}$ for some $\pi \in [0,1]$, then it also recommends $\{x_0\}$ for any $\pi' \in (\pi,1]$.

Theorem 1 (Cutoff strategies)**.** *Let $\alpha^* : [0,1] \to A$ be an optimal stationary strategy. Then, α^* is a cutoff strategy.*

By virtue of this theorem, we shall focus on stationary cutoff strategies in the sequel.

3 Single-Choice Benchmark

The problem that our DM faces departs from standard multi-armed bandit problems in two ways. First, x_0 and x_1 are negatively correlated: A success for one tool is conclusive evidence that the other is unproductive. Second, the DM is not restricted to choosing at most a single tool at a time; hence, she faces a *multi-choice* multi-armed bandit problem. This multi-choice feature of the setting allows the DM to accumulate more data by experimenting with both tools simultaneously, while the correlation feature allows her to *learn* about both tools from any single one. This section analyzes the single-choice benchmark, to isolate the second feature.

The set of allocations here is $A^0 := \{\alpha \in [0,1]^2 : \alpha_0 + \alpha_1 \leq 1\}$, and the Bellman equation becomes:

$$w(\pi) = \max_{\alpha \in A^0} \left\{ \alpha_0 \left[\lambda^0 \pi - c + \frac{\lambda^0 \pi \left(w\left(1\right) - w(\pi) \right) - \lambda^0 \pi (1-\pi) w'(\pi)}{\rho} \right] \right.$$
$$\left. + \alpha_1 \left[\lambda^0 (1-\pi) - c + \frac{\lambda^0 (1-\pi) \left(w\left(0\right) - w(\pi) \right) + \lambda^0 \pi (1-\pi) w'(\pi)}{\rho} \right] \right\}.$$

The same argument behind Theorem 1 applies to the single-choice problem. Let us first look for a cutoff strategy $\alpha(\cdot; \lambda^0, c) : [0,1] \to A^0$ of the following form:

- There is some $\underline{\pi} \in (0,1)$ such that, for all $\pi \in [0, \underline{\pi})$, $\alpha(\pi; \lambda^0, c) = (0,1)$; in words, for sufficiently low beliefs that $\omega = 0$, focus on tool x_1.

- There is some $\overline{\pi} \in (0,1)$, $\overline{\pi} > \underline{\pi}$, such that, for all $\pi \in (\overline{\pi}, 1]$, $\alpha(\pi; \lambda^0, c) = (1,0)$; namely, for sufficiently high beliefs that $\omega = 0$, focus on tool x_0.

- For all $\pi \in (\underline{\pi}, \overline{\pi})$, $\alpha(\pi; \lambda^0, c) = (0,0)$; if there is sufficient uncertainty about the state of nature, don't bother experimenting.

It remains to specify the strategy at the cutoffs. We do this below, after identifying the corresponding value function.

Under such a strategy, on $(0, \underline{\pi})$, we have:

$$-\lambda^0 \pi(1-\pi) w'(\pi) + (\rho + \lambda^0(1-\pi)) w(\pi) = \lambda^0(1-\pi)(\rho + \lambda^0 - c) - \rho c. \tag{1}$$

This equation is similar to equation (1) in Keller and Rady (2010). The homogeneous part of the solution is $w^H(\pi) := \pi \psi(\pi)^{-\frac{\rho}{\lambda^0}}$, where $\psi(\pi) := \frac{1-\pi}{\pi}$. Notice that $w^{H\prime}(\pi) = \frac{1-\pi+(\rho/\lambda^0)}{\pi(1-\pi)} w^H(\pi)$. For the particular part of the solution, we guess and verify $w^P(\pi) = a(1-\pi) + b$. For this guess to be correct, we must have:

$$\lambda^0 \pi(1-\pi) a + (\rho + \lambda^0(1-\pi))(a(1-\pi) + b) = \lambda^0(1-\pi)(\rho + \lambda^0 - c) - \rho c,$$

which gives $b = -c$ and $a = \lambda^0$. Up to a constant of integration C_1, the solution to this differential equation is $w(\pi) = C_1 \pi \psi(\pi)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c$.

On $(\overline{\pi}, 1)$, we have:

$$\lambda^0 \pi(1-\pi) w'(\pi) + (\rho + \lambda^0 \pi) w(\pi) = \lambda^0 \pi(\rho + \lambda^0 - c) - \rho c. \tag{2}$$

This equation is almost identical to equation (1) in Keller and Rady (2010); up to a constant of integration C_0, the solution is $w(\pi) = C_0(1-\pi)\psi(\pi)^{\frac{\rho}{\lambda^0}} + \lambda^0 \pi - c$.

We identify a candidate for an optimal strategy by pinning down $\underline{\pi}$, $\overline{\pi}$, C_1, and C_0. We do so by means of the *value-matching* (VM) and *smooth-pasting* (SP) conditions.

Condition (VM). $w(\underline{\pi}) = 0 = w(\overline{\pi})$

Condition (SP). $w'(\underline{\pi}) = 0 = w'(\overline{\pi})$

The (VM) condition says that, at the cutoff $\underline{\pi}$, the DM must be indifferent between trying out tool x_1 and giving up altogether; similarly for $\overline{\pi}$ and tool x_0. The (SP) condition says that, at the cutoffs, the marginal value of learning from the corresponding tool must equal that of the "no learning" choice, which equals 0. Without these conditions, the DM would be giving up either "too soon" or "too late."

The first equality in (VM) is

$$C_1 \underline{\pi} \psi(\underline{\pi})^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\underline{\pi}) - c = 0,$$

which gives $C_1 = C_1(\underline{\pi}) := -\frac{\lambda^0(1-\underline{\pi})-c}{\underline{\pi}}\psi(\underline{\pi})^{\frac{\rho}{\lambda^0}}$. The second equality in (VM) gives $C_0 = C_0(\overline{\pi}) := -\frac{\lambda^0\overline{\pi}-c}{1-\overline{\pi}}\psi(\overline{\pi})^{-\frac{\rho}{\lambda^0}}$. From (SP), we find:

$$\underline{\pi} = \frac{\lambda^0 - c}{\lambda^0}\frac{\lambda^0+\rho}{\lambda^0+\rho-c};$$

this expression lies in $(0,1)$ provided that $\lambda^0 > c$. With this expression for $\underline{\pi}$, we can write $C_1(\underline{\pi})$ as $C_1(\underline{\pi}) = \frac{\lambda^0 c}{\lambda^0+\rho}\psi(\underline{\pi})^{\frac{\rho}{\lambda^0}}$. The second equality in (SP) gives $\overline{\pi} = 1-\underline{\pi} \in (0,1)$, which leads to $C_0(\overline{\pi}) = C_1(\underline{\pi})\psi(\underline{\pi})^{-2\frac{\rho}{\lambda^0}}$.

To have a strategy of the form described in the three bullet points, with an intermediate range of beliefs in which the DM chooses the empty set, we must have $\overline{\pi} > \underline{\pi}$ (namely, $\underline{\pi} < 1/2$). This inequality holds if and only if $\rho(2c-\lambda^0) > \lambda^0(\lambda^0-c)$, which is possible only if $\lambda^0 < 2c$. (Recall that we assume that $\lambda^0 > c$.)

I will say that the tools are *expensive* if $\lambda^0 < 2c$. Indeed, $\lambda^0 < 2c$ means that, in the absence of learning, the empty set is more profitable than the complete toolkit. In this case, the only rationale for choosing the complete toolkit is its information value; but this value is not high enough for an *impatient* DM, one with $\rho > \frac{\lambda^0(\lambda^0-c)}{2c-\lambda^0}$. Conversely, we have that $\rho(2c-\lambda^0) \leq \lambda^0(\lambda^0-c)$ either if tools are *cheap*, $\lambda^0 \geq 2c$, or if the DM is *patient*, namely, if $\rho \leq \frac{\lambda^0(\lambda^0-c)}{2c-\lambda^0}$.

Case 1 (Expensive tools and impatience). $\rho(2c-\lambda^0) > \lambda^0(\lambda^0-c)$

Case 2 (Cheap tools and/or patience). $\rho(2c-\lambda^0) \leq \lambda^0(\lambda^0-c)$

Figure 2 portrays the partition of the space of "objective parameters" $\{(\lambda^0, c) \in \mathbb{R}^2_+ : \lambda^0 > c\}$ — excluding the "subjective" parameter ρ — according to whether they identify cheap tools or a sufficiently patient DM. The two cases correspond to the cases of low, intermediate, and high *stakes* in Klein and Rady (2011).[5]

Putting all of these pieces together, we identify the following candidate solution to the Bellman equation under case 1:

$$w^0(\pi) = \begin{cases} \frac{\lambda^0 c}{\lambda^0+\rho}\pi\left(\frac{\psi(\pi)}{\psi(\underline{\pi})}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c & \pi \in [0,\underline{\pi}]; \\ 0 & \pi \in (\underline{\pi},\overline{\pi}); \\ \frac{\lambda^0 c}{\lambda^0+\rho}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\overline{\pi})}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi - c & \pi \in [\overline{\pi},1]. \end{cases} \tag{3}$$

This function is continuously differentiable, strictly decreasing on $[0,\underline{\pi}]$, and strictly increasing on $[\overline{\pi},1]$ (see Lemma A1 in the appendix). Figure 3a on page 12 shows the plot of w^0 for $\lambda^0 = 0.7$, $c = 0.5$, and $\rho \in \{1, 10\}$.

[5] I thank Sven Rady for bringing this to my attention. In fact, with the lump sum from an arrival normalized to 1, and with s denoting the flow payoff from the safe arm, the case of low stakes in Klein and Rady (2011) can be rewritten as $\rho(2s-\lambda^0) > \lambda^0(\lambda^0-s)$; this is exactly the inequality in case 1, with s playing the role of c.

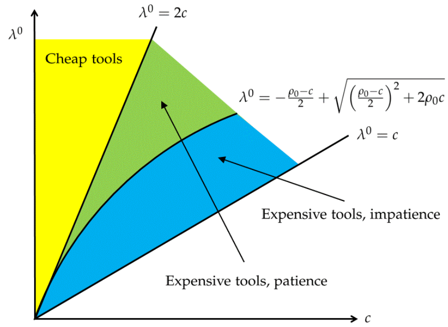


Figure 2: Objective-parameter space for a fixed $\rho = \rho_0$. The colored portion of the graph represents the parameter space for the problem. The yellow region represents cheap tools. The middle curve is the level curve of the threshold $\lambda^0(\lambda^0 - c) = \rho_0(2c - \lambda^0)$. In the green region, tools are expensive but $\rho = \rho_0$ is "patience enough." These two regions combined represent case 2; the blue region represents case 1.

Below cutoff $\underline{\pi}$, while the DM is experimenting on tool x_1 unsuccessfully, her posterior that the state is $\omega = 1$ gradually decreases. In between cutoffs, there is no experimentation, so beliefs remain frozen. Whether we recommend the DM to hold on to x_1 or to give up at $\underline{\pi}$, the result is always the same, well-defined dynamics for the posterior: If $\pi_0 < \underline{\pi}$, in the absence of successes, the posterior increases gradually until it reaches $\underline{\pi}$ and freezes there. The same is true for x_0 and $\overline{\pi}$: Whether the strategy recommends x_0 or to give up at $\overline{\pi}$, the path of posteriors starting from any prior in $(\overline{\pi}, 1]$ is well-defined. Thus, in the terminology of Klein and Rady (2011), all of these different specifications for the strategy at the thresholds are *admissible*.[6]

In this case, I will resolve indifference in favor of experimentation, and abuse terminology by talking about "the" optimal strategy. In the sequel, whenever possible, I will resolve indifference in favor of the largest bundle.

If the DM is sufficiently patient or the tools are sufficiently cheap — that is, under case 2 —, she may be willing to hold on to the singletons for longer. Consider a cutoff strategy $\alpha(\cdot; \lambda^0, c) : [0,1] \to A^0$ of the following form:

- For all $\pi \in [0, 1/2)$, $\alpha(\pi; \lambda^0, c) = (0,1)$; if $\pi < 1 - \pi$, that is, if state $\omega = 0$ is assessed to be less likely than state $\omega = 1$, focus on tool x_1.
- For all $\pi \in (1/2, 1]$, $\alpha(\pi; \lambda^0, c) = (1,0)$; if state $\omega = 0$ is believed to be more likely than state $\omega = 1$, focus on tool x_0 instead.

The recommendation at the cutoff $1/2$ is discussed below.

The (VM) condition is now $C_1 + \lambda^0 - 2c = C_0 + \lambda^0 - 2c$, which gives $C_1 = C_0 =: C$. The value

[6] Klein and Rady (2011) define a strategy to be admissible if, starting from any prior, the strategy yields a well-defined path of posteriors $t \mapsto \pi_t$.

of C is determined by (SP), $w'(1/2) = 0$, to be $C = \frac{(\lambda^0)^2}{\lambda^0+2\rho}$. Thus,

$$w^0\left(\frac{1}{2}\right) = \frac{\lambda^0(\rho + \lambda^0 - c) - 2\rho c}{\lambda^0 + 2\rho},$$

which is non-negative provided that $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$ (case 2).

We now have:

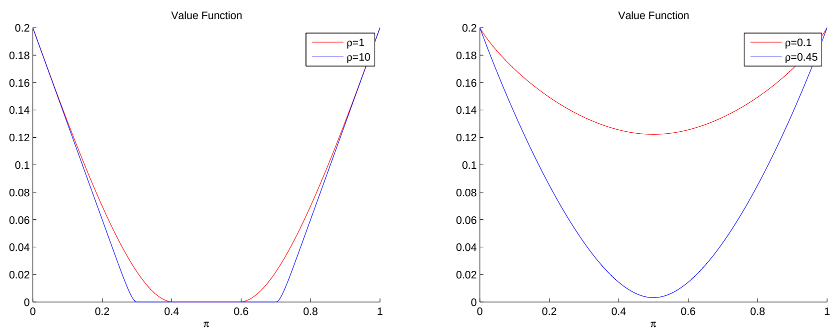
$$w^0(\pi) = \begin{cases} \frac{(\lambda^0)^2}{\lambda^0+2\rho}\pi\psi(\pi)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c & \pi \in \left[0, \frac{1}{2}\right]; \\ \frac{(\lambda^0)^2}{\lambda^0+2\rho}(1-\pi)\psi(\pi)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi - c & \pi \in \left(\frac{1}{2}, 1\right]. \end{cases} \tag{4}$$

Again, this function is continuously differentiable, strictly decreasing on $[0, 1/2)$, and strictly increasing on $(1/2, 1]$.[7] Figure 3b shows the plot of w^0 for $\lambda^0 = 0.7$, $c = 0.5$, and $\rho \in \{0.1, 0.45\}$.

To the right of the cutoff $1/2$, the DM experiments on tool x_1; while unsuccessful, her posterior gradually increases. To the left of $1/2$, instead, the posterior gradually decreases. If we specify that the DM should choose x_1 at the cutoff, we run into the following problem: The posterior is strictly increasing at $1/2$, but it switches sign and becomes strictly decreasing above $1/2$. This specification yields an inadmissible strategy; around the cutoff, the path of the posterior "chatters" (Romer, 1986). The same problem arises if we recommend x_0 at the cutoff instead.[8]

To obtain an admissible strategy, beliefs must freeze at the cutoff $1/2$. In the knife-edge case where the weak inequality in case 2 holds with equality, this can be achieved by specifying that the DM chooses the empty set at the threshold. Outside this knife-edge case, experimentation is strictly profitable; a way to prevent the posterior from chattering is to recommend the DM to *split* her time between x_0 and x_1. By dividing the intensity of experimentation equally between the two tools, beliefs (virtually) freeze at the threshold unless and until an arrival is observed.

The next theorem presents the optimal strategy in this single-choice benchmark.



(a) Graph of w^0 in (3) (b) Graph of w^0 in (4)

Figure 3: Graph of the value function in the single-choice benchmark.

[7] The proof of this assertion is very similar to that for w^0 in (3), Lemma A1; thus, any further details are omitted.

[8] I thank Sven Rady for making me aware of the presence of the admissibility problem in this case.

Theorem 2 (Single-choice benchmark). *Assume that $\lambda^0 > c$. Under case 1 — if $\rho(2c - \lambda^0) > \lambda^0(\lambda^0 - c)$ —, the cutoff strategy given by*

$$\alpha^\beta(\pi; \lambda^0, c) = \begin{cases} (0,1) & \pi \in \left[0, \underline{\pi}^\beta\right], \\ (0,0) & \pi \in \left(\underline{\pi}^\beta, \overline{\pi}^\beta\right), \\ (1,0) & \pi \in \left[\overline{\pi}^\beta, 1\right], \end{cases} \tag{5}$$

where $\underline{\pi}^\beta := \frac{\lambda^0 - c}{\lambda^0} \frac{\lambda^0 + \rho}{\lambda^0 + \rho - c} \in \left(0, \frac{1}{2}\right)$ and $\overline{\pi}^\beta := 1 - \underline{\pi}^\beta$, is the optimal strategy.[9] *Under case 2 — when $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$ —, the optimal strategy is given by:*

$$\alpha^\beta(\pi; \lambda^0, c) = \begin{cases} (0,1) & \pi \in \left[0, \frac{1}{2}\right), \\ \left(\frac{1}{2}, \frac{1}{2}\right) & \pi = \frac{1}{2}, \\ (1,0) & \pi \in \left(\frac{1}{2}, 1\right]. \end{cases} \tag{6}$$

When $\lambda^0 < 2c$, tools are costly for a DM who is sufficiently unsure about the state of nature. This cost outweighs the information value to an impatient DM, and she gives up. Instead, a sufficiently patient DM values this information, and never gives up. In either case, the DM hires the most promising tool — if it is sufficiently promising under impatience.

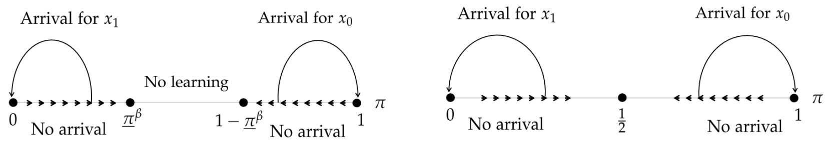
The experimentation dynamics under the optimal strategy in Theorem 2 are depicted in Figure 4. A sufficiently impatient DM who is unsure about the state of nature, one whose prior falls in the mid range $\left(\underline{\pi}^\beta, \overline{\pi}^\beta\right)$, "gives up" if tools are expensive; learning is simply too costly. If she is sufficiently confident about the state being $\omega = 1$ — namely, if her prior is in the low range $\left[0, \underline{\pi}^\beta\right]$ —, then the DM starts by renting tool x_1 alone. If this tool proves successful, then it is kept forever thereupon. While no arrivals occur for this tool, the DM revises her initial confidence, and becomes more and more pessimistic about this tool being productive. Of course, she becomes more and more optimistic about the state being $\omega = 0$. However, her posterior does not reach the point of being "optimistic enough" to switch to x_0: Eventually, her lost confidence leads her to drop x_1 and give up altogether, never giving tool x_0 a chance.

Unlike her more impatient counterpart, a sufficiently patient DM — or one facing cheap tools — never gives up. She starts with the tool about which she feels more confident, and holds on to it. Eventually, if there are no arrivals, her posterior approaches 1/2. At this point, she divides her attention or effort equally between the two tools until an arrival occurs.

Thus, an impatient DM tries at most a single tool, and never switches from one tool to the other. If she starts on tool x_1, she gives up after time

$$T_1 := \frac{\ln\left(\psi(\pi_0)\right) - \ln\left(\psi\left(\underline{\pi}^\beta\right)\right)}{\lambda^0}$$

[9] The superscript β is a mnemonic for "(single-choice) benchmark."



(a) Belief dynamics under strategy (5) (b) Belief dynamics under strategy (6)

Figure 4: Belief dynamics under the strategy identified in Theorem 2.

if no successes are observed; for x_0, she waits until

$$T_0 := \frac{\ln\left(\psi\left(\overline{\pi}^\beta\right)\right) - \ln\left(\psi(\pi_0)\right)}{\lambda^0}.$$

Finally, under case 2, the DM waits until

$$T' := \frac{\left|\ln\left(\psi\left(\frac{1}{2}\right)\right) - \ln\left(\psi(\pi_0)\right)\right|}{\lambda^0}.$$

4 Optimal Multi-Choice Strategy

In section 3, the DM can choose no more than one tool at a time. At most, she can divide her attention or effort between the two; but she cannot test them simultaneously. In this section, we allow her to try out both tools at once.

By Theorem 1, we focus on cutoff strategies. The complete toolkit yields better data, and information is more valuable when the DM is sufficiently unsure about the true state. Consider first a cutoff strategy $\alpha^*(\cdot;\lambda^0,c) : [0,1] \to A$ with the following properties:

- There is some $\underline{\pi} \in (0,1)$ such that, for all $\pi \in [0,\underline{\pi})$, $\alpha^*(\pi;\lambda^0,c) = (0,1,0)$; for sufficiently low beliefs that $\omega = 0$, focus on tool x_1.
- There is some $\overline{\pi} \in (0,1)$, $\overline{\pi} \geq \underline{\pi}$, such that, for all $\pi \in (\overline{\pi},1]$, $\alpha^*(\pi;\lambda^0,c) = (1,0,0)$; if state $\omega = 0$ is deemed sufficiently likely, then focus on tool x_0.
- For all $\pi \in (\underline{\pi},\overline{\pi})$, $\alpha^*(\pi;\lambda^0,c) = (0,0,1)$; if there is sufficient uncertainty about the state, then experiment on both tools simultaneously.

Under such a strategy, on $(0,\underline{\pi})$, we have the same differential equation as in section 3, equation (1). Thus, the solution has the same structure as before. The same is true on $(\overline{\pi},1)$, where equation (2) applies. On $(\underline{\pi},\overline{\pi})$, we now have

$$(\lambda^0+\rho)w(\pi) = \rho(\lambda^0-2c) + \lambda^0(\lambda^0-c);$$

solve for $w(\pi)$ to get $w(\pi) = \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho}$.

The (SP) conditions are the same as before; while the DM may enjoy a positive expected payoff when experimenting simultaneously on both tools, the marginal value of information is 0 (the payoff is constant) in this range. The (VM) conditions are now:

Condition (VM). $w(\underline{\pi}) = \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho} = w(\overline{\pi})$

The first equality in Condition (VM) is:

$$C_1 \underline{\pi} \psi(\underline{\pi})^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\underline{\pi}) - c = \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho},$$

which gives

$$C_1 = C_1(\underline{\pi}) := \frac{\lambda^0(\lambda^0+\rho)\underline{\pi} - \rho c}{\underline{\pi}(\lambda^0+\rho)} \psi(\underline{\pi})^{\frac{\rho}{\lambda^0}};$$

the second equality is very similar to the first, and gives

$$C_0 = C_0(\overline{\pi}) := \frac{\lambda^0(\lambda^0+\rho)(1-\overline{\pi}) - \rho c}{(1-\overline{\pi})(\lambda^0+\rho)} \psi(\overline{\pi})^{-\frac{\rho}{\lambda^0}}.$$

With the modified (VM) conditions, the first equality in Condition (SP) leads to:

$$\underline{\pi} = \frac{\lambda^0+\rho}{\lambda^0+\rho+c} \frac{c}{\lambda^0} \in (0,1).$$

With this expression for $\underline{\pi}$, we can write $C_1(\underline{\pi})$ as $C_1(\underline{\pi}) = \frac{\lambda^0 c}{\lambda^0+\rho} \psi(\underline{\pi})^{\frac{\rho}{\lambda^0}+1}$. The second equality in Condition (SP) is analogous to the first, and solving for $\overline{\pi}$ gives:

$$\overline{\pi} = \frac{(\lambda^0+\rho)(\lambda^0-c)+\lambda^0 c}{\lambda^0(\lambda^0+\rho+c)} = 1 - \underline{\pi} \in (0,1),$$

so $C_0(\overline{\pi}) = C_1(\underline{\pi})\psi(\underline{\pi})^{-2\left(\frac{\rho}{\lambda^0}+1\right)}$. Notice that $\overline{\pi} \geq \underline{\pi}$ (or $\underline{\pi} \leq 1/2$) if and only if $\rho(2c-\lambda^0) \leq \lambda^0(\lambda^0-c)$, namely, under case 2: The DM is willing to hold both tools when these are cheap or when she is patient enough to appreciate the additional information provided to her by simultaneous experimentation.

The candidate to solve the Bellman equation is now:

$$w^0(\pi) = \begin{cases} \frac{\lambda^0 c \psi(\underline{\pi})}{\lambda^0+\rho} \pi \left(\frac{\psi(\pi)}{\psi(\underline{\pi})}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c & \pi \in [0, \underline{\pi}); \\ \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho} & \pi \in [\underline{\pi}, \overline{\pi}]; \\ \frac{\lambda^0 c}{(\lambda^0+\rho)\psi(\overline{\pi})}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\overline{\pi})}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0 \pi - c & \pi \in (\overline{\pi}, 1]. \end{cases} \tag{7}$$

As in section 3, w^0 in (7) is continuously differentiable, strictly decreasing on $[0, \underline{\pi})$, and strictly increasing on $(\overline{\pi}, 1]$. Figure 5 shows the plot of w^0 for $\lambda^0 = 0.7$, $c = 0.3$, and $\rho \in \{0.1, 10\}$.

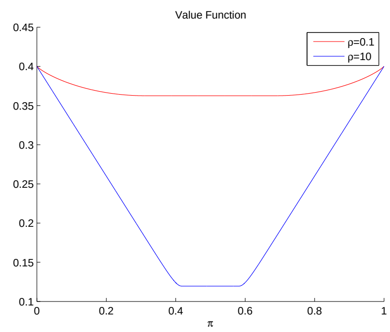


Figure 5: Graph of w^0 in (7); $\lambda^0 = 0.7$, $c = 0.3$, and $\rho \in \{0.1, 10\}$.

By choosing the complete toolkit, the DM can gather more data. Doing so allows her to distinguish between the events of failure of arrival altogether and arrival from a neglected tool, thus learning "faster." Notwithstanding, due to the negative correlation, any one tool provides some information about the other. The latter source of information is particularly relevant when tools are expensive; in this case, while the DM has the option to rent both tools at once, these may be too costly for her to exercise this option.

The next theorem presents the optimal strategy for the DM when she is allowed to engage in simultaneous experimentation. This result is the counterpart of Propositions 1 and 2 in Klein and Rady (2011).

Theorem 3 (Optimal strategy). *Under case 2 — if $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$ —, the cutoff strategy given by*

$$\alpha^*(\pi; \lambda^0, c) = \begin{cases} (0,1,0) & \pi \in [0, \underline{\pi}^*), \\ (0,0,1) & \pi \in [\underline{\pi}^*, \overline{\pi}^*], \\ (1,0,0) & \pi \in (\overline{\pi}^*, 1], \end{cases} \tag{8}$$

where $\underline{\pi}^* := \frac{\lambda^0+\rho}{\lambda^0+\rho+c}\frac{c}{\lambda^0} \in \left(0, \frac{1}{2}\right)$ and $\overline{\pi}^* := 1 - \underline{\pi}^*$, is the optimal strategy. Otherwise, under case 1, the optimal strategy is the same as in (5) *in Theorem 2.*[10]

Figure 6 portrays the dynamics of experimentation under the optimal strategy in Theorem 3. If her prior falls in the mid range $[\underline{\pi}^*, \overline{\pi}^*]$, the DM starts by renting both tools when these are cheap or when she is sufficiently patient. There is no updating while there are no arrivals. As soon as the first arrival occurs, the posterior jumps to either 1 or 0, depending on which one of the tools proves successful. Thenceforth, only the successful tool is kept.

Outside this range, the dynamics are similar to those under the strategy in Theorem 2. The difference is that, if the DM starts with a singleton and becomes gradually pessimistic, she will retain the first tool while eventually hiring the second one. Bad news about the chosen tool is

[10]We represent actions in A^0 as actions in A by adding a 0 as a third coordinate.

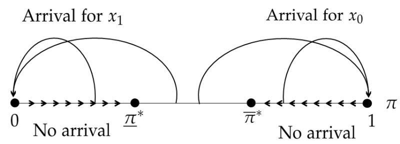


Figure 6: Belief dynamics under the strategy in (8)

good news about the other tool; but the lack of success leaves her unsure, rather than sufficiently confident about the neglected tool. Now, when unsure, the DM chooses both tools at once instead of, in the limit, dividing her attention between them. At this point, it may well happen that the first tool proves its worth after all, and the newly acquired tool is dismissed. If the new tool proves worthwhile instead, the first one is dismissed.

When the DM is sufficiently impatient, she does not take "advantage" of having the option to choose the complete toolkit if the tools are expensive. Otherwise, she never discards a tool unless and until the model uncertainty is fully resolved.

5 Extensions

This section analyzes several extensions to the basic model, or variations of it. The first subsection considers the case of asymmetric costs. The second discusses the case when choices are restricted to be nested. The third considers the case of "imperfect monitoring," namely, when the individual source of successes from employing both tools at once cannot be identified. The fourth and final subsection introduces a third, separate tool that the DM can try out before employing (any subset of) the original two.

5.1 Asymmetric Costs

This subsection discusses the case where the tools have different costs. Let $c_0, c_1 > 0$ denote the rental rates of x_0, x_1, respectively; assume that $c_1 < c_0 < \lambda^0$. Given $\pi \in [0,1]$, expected instantaneous rewards are now 0 from choosing the empty set; $\lambda^0\pi - c_0$ from $\{x_0\}$; $\lambda^0(1-\pi) - c_1$ from $\{x_1\}$; and $\lambda^0 - c_0 - c_1$ from X.

If the DM can only choose one tool at a time, as in section 3, her problem is represented by the following Bellman equation:

$$w(\pi) = \max_{\alpha \in A^0} \left\{ \alpha_0 \left[\lambda^0\pi - c_0 + \frac{\lambda^0\pi\left(w\left(1\right) - w(\pi)\right) - \lambda^0\pi(1-\pi)w'(\pi)}{\rho} \right] \right.$$
$$\left. + \alpha_1 \left[\lambda^0(1-\pi) - c_1 + \frac{\lambda^0(1-\pi)\left(w\left(0\right) - w(\pi)\right) + \lambda^0\pi(1-\pi)w'(\pi)}{\rho} \right] \right\}.$$

By analogous value-matching and smooth-pasting conditions to those in Section 3, we now get the following cutoffs for a strategy in which the DM gives up for some mid range of beliefs:

$$\underline{\pi}^\beta = \frac{\lambda^0+\rho}{\lambda^0}\frac{\lambda^0-c_1}{\lambda^0+\rho-c_1}, \quad \overline{\pi}^\beta = \frac{\rho c_0}{\lambda^0(\lambda^0+\rho-c_0)}.$$

We have that $\overline{\pi}^\beta > \underline{\pi}^\beta$ if and only if:

$$\lambda^0 < \frac{\rho}{\lambda+\rho}\frac{\lambda^0+\rho-c_1}{\lambda^0+\rho-c_0}c_0+c_1;$$

this condition reduces to $\lambda^0(\lambda^0-c)<\rho(2c-\lambda^0)$ (case 1) when $c_0=c_1=:c$, and it is satisfied if the costs are sufficiently high and the DM is sufficiently impatient. The counterpart of (3) is:

$$w^0(\pi) = \begin{cases} \frac{\lambda^0 c_1}{\lambda^0+\rho}\pi\left(\frac{\psi(\pi)}{\psi(\underline{\pi}^\beta)}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi)-c_1 & \pi\in\left[0,\underline{\pi}^\beta\right); \\ 0 & \pi\in\left[\underline{\pi}^\beta,\overline{\pi}^\beta\right]; \\ \frac{\lambda^0 c}{\lambda^0+\rho}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\overline{\pi}^\beta)}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi-c_0 & \pi\in\left(\overline{\pi}^\beta,1\right]. \end{cases} \tag{9}$$

Because of the more attractive immediate rewards, the DM experiments with the cheaper tool for a larger interval of beliefs: We have that $\underline{\pi}^\beta > 1-\overline{\pi}^\beta$ if and only if $c_1<c_0$. In the range of low beliefs, the DM tries tool x_1 out. The cutoff $\underline{\pi}^\beta$ is strictly decreasing in c_1; thus, as this tool becomes more expensive, the DM gives up "faster," or uses it for a smaller range of beliefs. Analogously, $\overline{\pi}^\beta$ is strictly increasing in c_0; as x_0 becomes more expensive, the DM requires higher confidence in it to employ it.

The case of cheap tools or sufficiently patient DM corresponds now to the case:

$$\lambda^0 \geq \frac{\rho}{\lambda+\rho}\frac{\lambda^0+\rho-c_1}{\lambda^0+\rho-c_0}c_0+c_1.$$

The smooth-pasting condition is the same as before: If π_0^β denotes the threshold, we have $w'\left(\pi_0^\beta\right)=0$. The value-matching condition identifies the threshold as the higher root of the following quadratic equation:

$$(c_0-c_1)(\lambda^0)^2x^2-\left[(c_0-c_1)(\lambda^0)^2-2\rho\lambda^0(\lambda^0+\rho)\right]x-\rho(\lambda^0+\rho)(\lambda^0+c_0-c_1)=0;$$

this equation yields $x=1/2$ as unique solution if $c_0=c_1$. Again, the cheapest tool is the one employed for longer: $\pi_0^\beta>1/2$. The counterpart of (4) is:

$$w^0(\pi) = \begin{cases} \frac{(\lambda^0)^2\left(1-\pi_0^\beta\right)}{\lambda^0\left(1-\pi_0^\beta\right)+\rho}\pi\left(\frac{\psi(\pi)}{\psi(\pi_0^\beta)}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi)-c_1 & \pi\in\left[0,\pi_0^\beta\right]; \\ \frac{(\lambda^0)^2\pi_0^\beta}{\lambda^0\pi_0^\beta+\rho}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\pi_0^\beta)}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi-c_0 & \pi\in\left(\pi_0^\beta,1\right]. \end{cases} \tag{10}$$

The function in (10) is non-negative provided that

$$\pi_0^\beta \geq \frac{\rho c_0}{\lambda^0(\lambda^0 + \rho - c_0)} = \overline{\pi}^\beta.$$

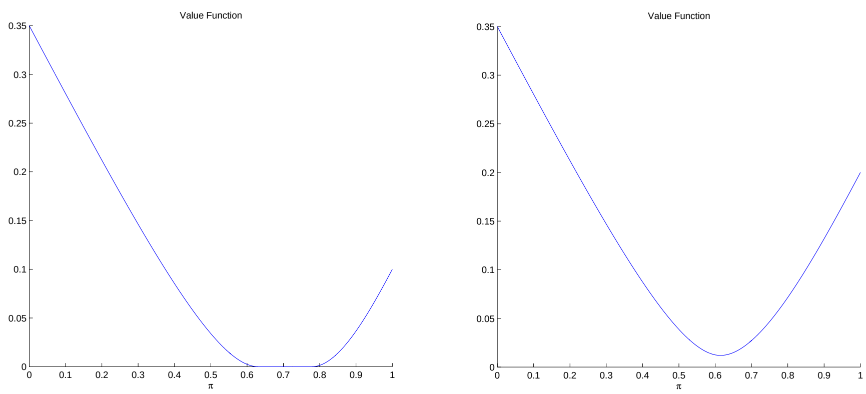
This inequality says that the DM should not experiment on $\{x_0\}$ for a larger range of beliefs than she would when she entertains the possibility of giving up. Under symmetric costs, this inequality is automatically satisfied in the range of parameters for which the corresponding strategy is optimal; the DM holds on to a failing $\{x_0\}$ until experimentation is no longer worth her while. But here, she has in $\{x_1\}$ a cheaper, more attractive alternative that allows her to continue learning.

Figure 7a shows the plot of w^0 in (9) in the case $(c_0, c_1) = (0.6, 0.35)$, $\lambda^0 = 0.7$, and $\rho = 0.99$. Figure 7b depicts w^0 in (10), in the case $(c_0, c_1) = (0.5, 0.35)$, $\lambda^0 = 0.7$, and $\rho = 0.99$.

The portion of the single-choice strategy corresponding to case 1, under which the DM gives up for an intermediate range of beliefs, is also part of the optimal strategy under simultaneous experimentation. Therefore, I describe this instance of the strategy below, in the simultaneous-experimentation case. As for the strategy that gives rise to value function (10), the issue of admissibility applies. At the cutoff π_0^β, the DM splits her time between the two tools.

If we allow the DM to choose both tools at once, the Bellman equation of the problem is:

$$\begin{aligned} w(\pi) = \max_{\alpha \in A} \Bigg\{ & \alpha_0 \left[\lambda^0 \pi - c_0 + \frac{\lambda^0 \pi \left[w\left(1\right) - w(\pi) \right] - \lambda^0 \pi (1-\pi) w'(\pi)}{\rho} \right] \\ & + \alpha_1 \left[\lambda^0 (1-\pi) - c_1 + \frac{\lambda^0 (1-\pi) \left[w\left(0\right) - w(\pi) \right] + \lambda^0 \pi (1-\pi) w'(\pi)}{\rho} \right] \\ & + \alpha_2 \left[\lambda^0 - c_0 - c_1 + \frac{\lambda^0 \pi \left[w\left(1\right) - w(\pi) \right] + \lambda^0 (1-\pi) \left[w\left(0\right) - w(\pi) \right]}{\rho} \right] \Bigg\}. \end{aligned}$$



(a) Graph of w^0 in (9); $c_0 = 0.6$, $c_1 = 0.35$. (b) Graph of w^0 in (10); $c_0 = 0.5$, $c_1 = 0.35$

Figure 7: Value function under asymmetric costs, given $\lambda^0 = 0.7$ and $\rho = 0.99$.

The conditions to identify the optimal strategy having the form in section 4 when the DM experiments with both tools at once are:

Condition (VM-a). $w(\underline{\pi}) = \lambda^0 - c_1 - \frac{\rho c_0}{\lambda^0+\rho} + \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}\underline{\pi}$

Condition (VM-b). $w(\overline{\pi}) = \lambda^0 - c_1 - \frac{\rho c_0}{\lambda^0+\rho} + \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}\overline{\pi}$

Condition (SP-a). $w'(\underline{\pi}) = \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}$

Condition (SP-b). $w'(\overline{\pi}) = \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}$

From these conditions, we get:

$$\underline{\pi}^* = \frac{\lambda^0+\rho}{\lambda^0+\rho+c_1}\frac{c_0}{\lambda^0}, \quad \overline{\pi}^* = \frac{(\lambda^0+\rho)(\lambda^0-c_1)+\lambda^0 c_0}{\lambda^0(\lambda^0+\rho+c_0)}.$$

We have that $\overline{\pi}^* \geq \underline{\pi}^*$ if and only if:

$$\lambda^0(\lambda^0-c_1)+\lambda^0 c_0\left[\frac{c_1-c_0}{\lambda^0+\rho+c_1}\right] \geq \rho\left[\frac{\lambda^0+\rho+c_0}{\lambda^0+\rho+c_1}c_0+c_1-\lambda^0\right].$$

This condition reduces to case 2, $\lambda^0(\lambda^0-c) \geq \rho(2c-\lambda^0)$, when $c_0 = c_1 =: c$, and holds if the costs are sufficiently low or if the DM is sufficiently patient. The counterpart of (7) is:

$$w^0(\pi) = \begin{cases} \frac{\lambda^0 c_0 \psi(\underline{\pi})}{\lambda^0+\rho}\pi\left(\frac{\psi(\pi)}{\psi(\underline{\pi})}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c_1 & \pi \in [0, \underline{\pi}^*); \\ \lambda^0 - c_1 - \frac{\rho c_0}{\lambda^0+\rho} + \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}\pi & \pi \in [\underline{\pi}^*, \overline{\pi}^*]; \\ \frac{\lambda^0 c_1}{(\lambda^0+\rho)\psi(\overline{\pi})}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\overline{\pi})}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi - c_0 & \pi \in (\overline{\pi}^*, 1]. \end{cases} \quad (11)$$

Figure 8 shows the plot of w^0 for the case $(c_0, c_1) = (0.5, 0.35)$, $\lambda^0 = 0.7$, and $\rho = 0.99$.

This function is non-negative provided that its minimizer, which falls in the range $(\overline{\pi}^*, 1]$, is at least at large as $\frac{\rho c_0}{\lambda^0(\lambda^0+\rho-c_0)} = \overline{\pi}^\beta$. This condition says that the range of beliefs at which the marginal value corresponding to $\{x_0\}$ is positive cannot be larger than the corresponding range under the single-choice benchmark. In the latter case, when there is no experimentation for beliefs in the mid range, the value function is strictly increasing exactly in the range in which $\{x_0\}$ is recommended. Here, however, the complete set is employed for mid-range beliefs; and with $c_0 > c_1$, the slope of the value function is negative on this range and becomes positive after the minimizer, after having switched to $\{x_0\}$. For optimality, we want that this further descent after having switched to $\{x_0\}$ does not dip the value function below 0.

Under symmetric costs, the case $\overline{\pi}^* < \underline{\pi}^*$ corresponds to the case $\overline{\pi}^\beta > \underline{\pi}^\beta$. This need not be the case under asymmetric costs. A sufficient condition for such correspondence, which allows

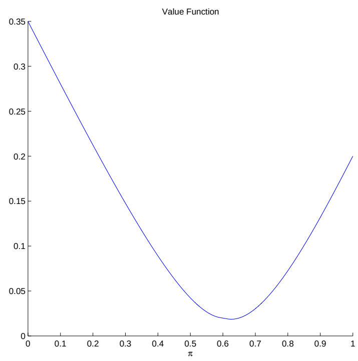


Figure 8: Graph of w^0 in (11); $\lambda^0 = 0.7$, $\rho = 0.99$, $c_0 = 0.5$, and $c_1 = 0.35$.

the counterpart of the strategy in section 4 to be optimal under asymmetric costs, is the following:

$$c_0 > \lambda^0 \frac{\lambda^0 + \rho}{\lambda^0 + 2\rho}; \quad c_1 > \frac{\lambda^0 + \rho}{\rho}(\lambda^0 - c_0).$$

While both costs must be lower than λ^0, they should not be "too low."

Theorem 4 (Asymmetric costs)**.** *Assume that* $\lambda^0 > c_0 > c_1$; $c_0 > \lambda^0 \frac{\lambda^0+\rho}{\lambda^0+2\rho}$; $c_1 > \frac{\lambda^0+\rho}{\rho}(\lambda^0 - c_0)$. *If*

$$\lambda^0(\lambda^0 - c_1) + \lambda^0 c_0 \left[\frac{c_1 - c_0}{\lambda^0 + \rho + c_1}\right] \geq \rho \left[\frac{\lambda^0 + \rho + c_0}{\lambda^0 + \rho + c_1} c_0 + c_1 - \lambda^0\right],$$

and if the minimum of (11) *is at least as high as:*

$$\frac{(\lambda^0 + \rho)(\lambda^0 - c_1)(\lambda^0 + \rho + c_1) - \lambda^0 c_0 (c_0 - c_1) - \rho c_0 (\lambda^0 + \rho + c_0)}{(2\rho + \lambda^0)(\lambda^0 + \rho + c_0)} \geq 0,$$

the strategy given by (8) *with* $\underline{\pi}^* = \frac{\lambda^0+\rho}{\lambda^0+\rho+c_1}\frac{c_0}{\lambda^0}$ *and* $\overline{\pi}^* = \frac{(\lambda^0+\rho)(\lambda^0-c_1)+\lambda^0 c_0}{\lambda^0(\lambda^0+\rho+c_0)}$ *is the optimal strategy. If*

$$\lambda^0(\lambda^0 - c_1) + \lambda^0 c_0 \left[\frac{c_1 - c_0}{\lambda^0 + \rho + c_1}\right] < \rho \left[\frac{\lambda^0 + \rho + c_0}{\lambda^0 + \rho + c_1} c_0 + c_1 - \lambda^0\right],$$

the optimal strategy is the same as in (5)*, with* $\underline{\pi}^\beta = \frac{\lambda^0+\rho}{\lambda^0}\frac{\lambda^0-c_1}{\lambda^0+\rho-c_1}$ *and* $\overline{\pi}^\beta = \frac{\rho c_0}{\lambda^0(\lambda^0+\rho-c_0)}$.

The lower bound on the minimum, which is non-negative in the corresponding case of low costs or sufficient patience, helps handle the non-monotonicity of the value function on the range in which the optimal strategy recommends $\{x_0\}$. In the case where tools are expensive and the DM is sufficiently impatient, this lower bound is negative, while the corresponding value function has 0 as its minimum. In this case, the additional condition is redundant.[11]

[11]No such additional condition is needed under symmetric costs.

5.2 Nested Choices

So far, the DM has had the option to hire a previously ignored tool, and to "re-hire" a tool that has been previously tried out and dismissed. However, it may be that such tools "disappear." For instance, a neglected tool may become rusty; a research project that is set aside may be scooped by another researcher; an overlooked applicant or a dismissed employee may find another job and exit the market. In this subsection, I consider the extreme case where choices must be nested, so once a tool is ignored or discarded it can never be chosen. This restriction introduces an option value to holding on to tools beyond what a less-restricted DM would consider optimal.

Now, the state space keeps track not only of the beliefs of the DM, but also of her feasible set of choices. For simplicity of the discussion, I restrict the DM to spending all of each time interval on a single toolkit; that is, I consider the restricted action space $A^r := \{\alpha \in \{0,1\}^3 : \alpha_0 + \alpha_1 + \alpha_2 \leq 1\}$.[12]

Let $w^r : [0,1] \times 2^X \to \mathbb{R}$ represent the restricted value function. Clearly, $w^r(\pi, \emptyset) = 0$. The Bellman equation for $w^r(\pi, \{x_1\})$ is:

$$w^r(\pi,\{x_1\}) = \max\Big\{0, w^r(\pi,\{x_1\}) + \Big[\lambda^0(1-\pi) - c \\ + \frac{\lambda^0}{\rho}(1-\pi)\left(\lambda^0 - c - w^r(\pi,\{x_1\}) + \pi(1-\pi)w^{r\prime}(\pi,\{x_1\})\right) - w^r(\pi,\{x_1\})\Big]\rho \mathrm{d}t\Big\};$$

either $w^r(\pi,\{x_1\}) = 0$, or $w^r(\pi,\{x_1\})$ solves the same differential equation (1) as w does in Section 3. Looking for a cutoff strategy, the same (VM) and (SP) conditions relating the choice of $\{x_1\}$ and the choice of the empty set apply. Thus, we have:

$$w^r(\pi,\{x_1\}) = \begin{cases} \frac{\lambda^0 c}{\lambda^0+\rho}\pi\left(\frac{\psi(\pi)}{\psi(\underline{\pi}_1^r)}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c & \pi \in [0, \underline{\pi}_1^r], \\ 0 & \pi \in (\underline{\pi}_1^r, 1], \end{cases} \tag{12}$$

where $\underline{\pi}_1^r := \frac{\lambda^0 - c}{\lambda^0}\frac{\lambda^0+\rho}{\lambda^0+\rho-c} = \underline{\pi}^\beta$. The same argument applies to $w^r(\pi,\{x_0\})$, leading to:

$$w^r(\pi,\{x_0\}) = \begin{cases} 0 & \pi \in [0, \underline{\pi}_0^r), \\ \frac{\lambda^0 c}{\lambda^0+\rho}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\underline{\pi}_0^r)}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi - c & \pi \in [\underline{\pi}_0^r, 1], \end{cases} \tag{13}$$

where $\underline{\pi}_0^r := \overline{\pi}^\beta$. Finally, for $w^r(\pi, X)$, we have:

$$w^r(\pi, X) = \max\{w^r(\pi,\{x_0\}), w^r(\pi,\{x_1\}), \\ \left(\lambda^0 - 2c + \frac{\lambda^0(\lambda^0 - c - w^r(\pi,X))}{\rho} - w^r(\pi,X)\right)\mathrm{d}t + w^r(\pi,X)\Big\}.$$

As before, we look for an optimal cutoff strategy that recommends the singleton $\{x_1\}$ for sufficiently low π, and the singleton $\{x_0\}$ for sufficiently high π. Assume that $\rho(2c-\lambda^0) \leq \lambda^0(\lambda^0 - c)$;

[12] The superscript r is a mnemonic for "restricted."

hence, $\underline{\pi}^\beta \geq 1/2 \geq \overline{\pi}^\beta$. (The case $\underline{\pi}^\beta < \overline{\pi}^\beta$ is handled similarly.) On $[0, \overline{\pi}^\beta)$,

$$w^r(\pi, X) = \max\left\{w^r(\pi, \{x_1\}), \left(\lambda^0 - 2c + \frac{\lambda^0(\lambda^0 - c - w^r(\pi, X))}{\rho} - w^r(\pi, X)\right) \mathrm{d}t + w^r(\pi, X)\right\}.$$

We look for a cutoff $\underline{\pi} \in (0, \overline{\pi}^\beta)$ such that $w^r(\underline{\pi}, \{x_1\}) = \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho}$. Similarly, on $(\underline{\pi}^\beta, 1]$,

$$w^r(\pi, X) = \max\left\{w^r(\pi, \{x_0\}), \left(\lambda^0 - 2c + \frac{\lambda^0(\lambda^0 - c - w^r(\pi, X))}{\rho} - w^r(\pi, X)\right) \mathrm{d}t + w^r(\pi, X)\right\},$$

and we seek an analogous cutoff $\overline{\pi} \in (\underline{\pi}^\beta, 1)$ for $w^r(\pi, \{x_0\})$.

The existence of these cutoffs is established in Lemma A3 in the appendix. The remaining details of the optimal strategy are provided in the following theorem.

Theorem 5 (Nested choices). *There exist two unique cutoffs denoted by $\underline{\pi}^r \in (0, \min\{\overline{\pi}^\beta, \underline{\pi}^\beta\})$, $\overline{\pi}^r \in (\max\{\overline{\pi}^\beta, \underline{\pi}^\beta\}, 1)$ such that, when choices must be nested, the optimal strategy is as follows:*

- $\alpha^r(\pi, \emptyset; \lambda^0, c) = (0,0,0)$;
- $\alpha^r(\pi, \{x_1\}; \lambda^0, c) = \begin{cases} (0,1,0) & \pi \in \left[0, \underline{\pi}^\beta\right], \\ (0,0,0) & \pi \in \left(\underline{\pi}^\beta, 1\right]; \end{cases}$
- $\alpha^r(\pi, \{x_0\}; \lambda^0, c) = \begin{cases} (0,0,0) & \pi \in \left[0, \overline{\pi}^\beta\right), \\ (1,0,0) & \pi \in \left[\overline{\pi}^\beta, 1\right]; \end{cases}$
- *If $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$ (case 2), then* $\alpha^r(\pi, X; \lambda^0, c) = \begin{cases} (0,1,0) & \pi \in [0, \underline{\pi}^r), \\ (0,0,1) & \pi \in [\underline{\pi}^r, \overline{\pi}^r], \\ (1,0,0) & \pi \in (\overline{\pi}^r, 1]; \end{cases}$
- *If $\rho(2c - \lambda^0) > \lambda^0(\lambda^0 - c)$ (case 1), then $\alpha^r(\pi, X; \lambda^0, c)$ is as in* (5) *in Theorem 2.*

At the outset, the feasible set for the DM is all of X. If the tools are expensive, a sufficiently impatient DM behaves as her unrestricted counterpart does. However, in the presence of cheap tools or sufficient patience, the DM has to be more certain about the state to go with a singleton: If $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$, then $\underline{\pi}^r < \overline{\pi}^\beta \leq \underline{\pi}^* \leq \overline{\pi}^* \leq \underline{\pi}^\beta < \overline{\pi}^r$. Intuitively, by starting with a singleton, she is giving up the option value of being able to switch to either of the tools at a later point in time — after having gathered more information.

5.3 Imperfect Monitoring

In section 4, we have assumed that the DM can observe from which tool successes come when renting both tools at once. In other words, she can "monitor" both tools closely and identify the source of a success when they are simultaneously employed. In applications, it may be the case that the output of a team can only be measured with respect to the team, and the individual contributions of the team members cannot be readily assessed. This subsection considers the case where the DM, when renting both tools, can only observe the occurrence of arrivals but not their "precedence"; all she observes is whether the toolkit X has produced value.

Figure 9 describes ex-post payoffs to the DM and the data she now collects from the complete toolkit. In this variation of the problem, the complete toolkit is as uninformative about the state as the empty toolkit; the only difference between the two is that the former yields an immediate payoff of $\lambda^0 - 2c$. To learn about the state, the DM must give tools a chance to stand *on their own*. Unlike under the main specification, "experimentation" now entails focusing on singletons.

The next theorem presents the optimal strategy under this alternative specification.

Theorem 6 (Imperfect monitoring). *Assume that $\rho(\lambda^0 - 2c) > \lambda^0 c$. The cutoff strategy given by*

$$\alpha^*(\pi; \lambda^0, c) = \begin{cases} (0,1,0) & \pi \in [0, \underline{\pi}'), \\ (0,0,1) & \pi \in [\underline{\pi}', \overline{\pi}'], \\ (1,0,0) & \pi \in (\overline{\pi}', 1], \end{cases} \tag{14}$$

where $\underline{\pi}' := \frac{\lambda^0+\rho}{\rho+c}\frac{c}{\lambda^0} \in \left(0, \frac{1}{2}\right)$ and $\overline{\pi}' := 1 - \underline{\pi}'$, is the optimal strategy. If $\rho(\lambda^0 - 2c) \leq \lambda^0 c$, the optimal strategy is as in Theorem 2: It is given by (6) *if $-\lambda^0(\lambda^0 - c) \leq \rho(\lambda^0 - 2c)$ (case 2), and by* (5) *if $-\lambda^0(\lambda^0 - c) > \rho(\lambda^0 - 2c)$ (case 1).*

When $\lambda^0 < 2c$, tools are too costly for the DM to ever want to carry both of them at once. The same is true if they are cheap but the DM is sufficiently patient: While the full toolkit may be attractive, more so is the information that only singletons can provide. If the tools are sufficiently cheap, an impatient DM rents both and enjoys her constant expected payoff when she is sufficiently unsure about the state; not appreciating the additional information, she is not willing to give up the higher instant surplus to learn about the tools. Instead, a more patient

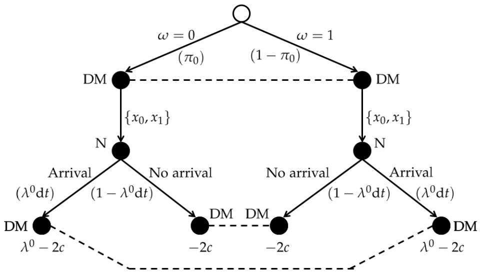


Figure 9: Observations and payoffs from choosing the complete toolkit, under imperfect monitoring.

DM is happy to stick to singletons for longer, due to their information value. Figure 10 is the counterpart of Figure 2 under this alternative specification.

The evolution of beliefs under the strategy in Theorem 6 is similar to that under Theorems 2 and 3; the main difference is that now there is no learning from the complete toolkit. Thus, once the DM chooses the full toolkit (optimally), she sticks to it thereupon.

5.4 Adding a Third Project

So far, the DM has only been given two projects or tools to choose from, each of which is equally appealing in its corresponding state. One and only one of them is fruitful, and the only problem is determining which one it is. In this section, we give the DM a third project she can work on, or a third tool she can employ. If this third tool is productive, its arrival rate is higher than that of the other two. But this new tool may be unproductive, while one of the original two tools is certainly productive; this new project may never flourish, while one of the original two eventually will. For simplicity, I assume that this third tool is "incompatible" with the other two in the sense that it requires the full attention of the DM while she is employing it; and that it must be forsaken once ignored or abandoned. Thus, the problem is to determine for how long to experiment with the new tool, if at all, before switching to the original toolkit.

There are now three tools, $X = \{x_0, x_1, y\}$. The DM allocates her time between the different subsets of $\{x_0, x_1\}$ and $\{y\}$. Employing this new tool also involves a cost of $c_y > 0$. There is a new state of nature, $\theta \in \{0, 1\}$, the realization of which is also unobserved; the new tool produces successes at a rate $\lambda^1\theta$, where $\lambda^1 > \lambda^0$ is *known*. Thus, the DM knows that, if this new tool proves successful, it is more appealing than any of the others; otherwise, she is better off with the original toolkit.

Let $\mu \in [0, 1]$ denote the assessment of the DM that $\theta = 1$; $\mu^0 \in [0, 1]$ denotes the corresponding prior. Representing separate tools, I assume that ω and θ are independent. Thus, since the

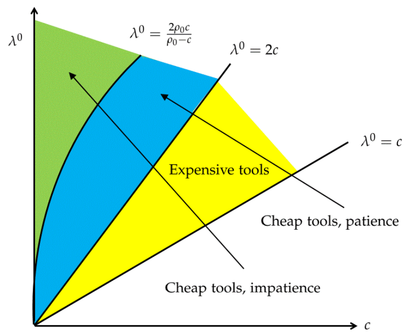


Figure 10: Objective-parameter space for a fixed $\rho = \rho_0$. The colored portion of the graph represents the parameter space for the problem. The yellow region represents expensive tools. The curve dividing the green and blue regions represents the level curve of the threshold $\rho_0(\lambda^0 - 2c) = \lambda^0 c$. In the green region, tools are cheap and the DM is impatient; in the blue region, the DM is patient instead.

DM updates her beliefs about each state from different data, the posteriors are also independent. In the event of an arrival from y, the posterior on θ jumps to 1; after spending a fraction $\gamma \in [0, 1]$ of time on y without observing any success, her posterior on θ gradually decreases according to $\dot{\mu} = -\gamma\lambda^1\mu(1-\mu)$. Immediate rewards are $\lambda^1\mu - c_y$.

Under irreversibility, once the DM switches away from y, she is back to the problem analyzed in previous sections; she divides her time among x_0, x_1, or both, according to the corresponding optimal strategy, and she enjoys a continuation payoff of $w^0(\pi)$ if π is the posterior that $\omega = 0$ at the time of switching. Thus, while she has not yet switched away from y, the new Bellman equation is:

$$w(\pi,\mu) = \max_{\alpha\in[0,1]} \left\{ \alpha w^0(\pi) + (1-\alpha)\left[\lambda^1\mu - c_y + \frac{\lambda^1\mu\left(w\left(\pi,1\right) - w(\pi,\mu)\right) - \lambda^1\mu(1-\mu)w_2'(\pi,\mu)}{\rho}\right]\right\}.$$

Consider a strategy such that, for each $\pi \in [0, 1]$, there is some $\mu(\pi) \in [0, 1]$ such that the DM starts employing tool y if $\mu \geq \mu(\pi)$, and follows the optimal strategy among $\{x_0, x_1\}$ otherwise. On the region of the state space where the DM experiments with y, we have:

$$w(\pi,\mu) = \lambda^1\mu - c_y + \frac{\lambda^1\mu\left(w\left(\pi,1\right) - w(\pi,\mu)\right) - \lambda^1\mu(1-\mu)w_2'(\pi,\mu)}{\rho}.$$

By assumption, $w\left(\pi, 1\right) = \lambda^1 - c_y$. Thus, this equation is analogous to equation (2), and yields:

$$w(\pi,\mu) = C(\pi)(1-\mu)\psi(\mu)^{\frac{\rho}{\lambda^1}} + \lambda^1\mu - c_y,$$

where $C(\cdot)$ is some continuously differentiable function. The value-matching condition gives $w(\pi, \mu(\pi)) = w^0(\pi)$ for all $\pi \in [0, 1]$; the smooth-pasting condition is $w_2'(\pi, \mu(\pi)) = 0$. Combining these two conditions gives:

$$\mu(\pi) = \frac{\rho}{\lambda^1}\frac{w^0(\pi) + c_y}{\lambda^1 + \rho - w^0(\pi) - c_y},$$

and $C(\pi) = \frac{\lambda^1}{\lambda^1+\rho}\psi(\mu(\pi))^{-\frac{\rho}{\lambda^1}}$.

Theorem 7 (Third tool)**.** *The optimal strategy consists of starting on $\{y\}$ provided that $\mu^0 \geq \mu(\pi^0)$, and sticking to it while the posterior (π, μ) satisfies $\mu \geq \mu(\pi)$; if $\mu < \mu(\pi)$, switch to the optimal strategy of Theorem 3 on the toolkit $\{x_0, x_1\}$.*

Figure 11 depicts the threshold $\mu(\pi)$ corresponding to the value function in Figure 5. If the DM is sufficiently sure about ω, she has to be sufficiently confident that $\theta = 1$ to start on tool y. If she is unsure about ω, she employs tool y for a wider range of beliefs on θ: If she switches to the original tools, she will rent both x_0 and x_1 at once and bear a high cost of experimentation. While experimenting unsuccessfully on y, her beliefs about $\theta = 1$ gradually decline. If a success

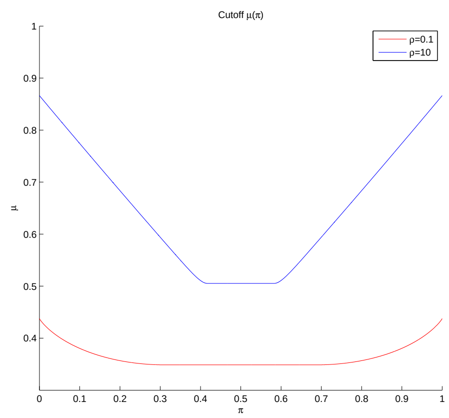


Figure 11: Cutoff $\mu(\pi)$ corresponding to $w^0(\pi)$ in Figure 5; $c_y = c = 0.3$. Above the cutoff, the optimal strategy recommends $\{y\}$; below the cutoff, it recommends following the optimal strategy for $\{x_0, x_1\}$.

occurs before she switches, she learns that $\theta = 1$, and sticks to y thereupon.

6 Conclusion

This paper analyzes the experimentation problem faced by a decision maker who can work on up to two problems at once over time. One and only one of these can produce successes, but the decision maker does not know ex-ante which one. To learn about the projects, she may work on one at a time, exploiting their correlation, or on both at once, gathering more data.

If experimentation is cheap, or if she is sufficiently patient, the decision maker starts by working on both projects at once if she is sufficiently uncertain about the state of nature; more so if projects ignored or discarded can be scooped. Working on both projects at once, she can identify the profitable one as soon as the first breakthrough occurs, learning nothing new in the meantime — lack of success on both projects is a "neutral" event.

If she is sufficiently sure about a project, she starts working on it exclusively. As long as she encounters no success, she becomes gradually pessimistic. Eventually, once she becomes sufficiently unsure, she takes on the other project as well. Since bad news about one project is good news about the other one, one might expect that the decision maker would switch from one to the other after meeting with no successes. However, the lack of success leaves her unsure about which project is better, rather than confident enough about the neglected project. At this point, she works on both projects at once until her uncertainty is resolved.

If she does not have the option to work on more than one project simultaneously, the decision maker works on individual projects for longer. If costs are high and the decision maker is sufficiently impatient, she eventually gives up if there are no arrivals. In this case, the neglected project is never given a chance: Before her posterior reaches a level where it would be optimal to work on it, it reaches a level of uncertainty such that neither project is worth pursuing any further. In fact, such a decision maker never works on both projects at once even if she can.

When the projects cannot be "individually monitored" unless they are studied in isolation, the only way for the decision maker to learn is to stick to singletons, and to test the alternatives "on their own." In this case, it is the impatient decision maker the one who works on both projects at once when experimentation is cheap; she is not willing to give up the higher instant surplus to learn about the state of nature.

The structure of the problem studied here is extremely simple. However, while some extensions are feasible, such as allowing for "interior" states $0 < \underline{\omega} < \overline{\omega} < 1$, the problem can become intractable or far too cumbersome very quickly. Allowing for a richer set of states would be interesting, but it requires expanding the dimensionality of the state space.[13] The limitations in this direction are substantial under the standard approach to control problems. Thus, for a richer discrete-time problem in a similar vein, Francetich and Kreps (2014) explores heuristics.

Another interesting extension of the problem studied in this paper is the extension to strategic experimentation. For example, imagine there are two agents experimenting simultaneously. Would they ever want to experiment on both projects at once, if they can learn from each other? If the alternatives are tools for rent from a hardware store, how should the store set the rental costs, or auction off the tools? In the example of the professional-services firm, the manager of the firm is a principal who learns about the dexterity or productivity of her employees. But these employees may be agents who exert unobservable effort in their work. For failures to produce breakthroughs to be informative of the state of nature, instead of being simply a reflection of shirking, wages and compensations must incentivize the agents to work hard. How much effort should the agents be induced to exert, and how should they be compensated for their effort? This is the subject of ongoing work.

References

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

Forand, J.G. (2013). Keeping your options open.

Francetich, A. and Kreps, D. (2014). Choosing a good toolkit: An essay in behavioral economics. Working Paper.

Gittins, J. and Jones, D. (1974). A dynamic allocation index for the sequential design of experiments. In: *Progress in Statistics*, pp. 241–266. Amsterdam: North Holland.

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

[13]Some work has been done dealing with vector-valued states, such as Klein and Rady (2011) and Forand (2013); however, their analysis is catered to their formulation, and does not readily extend to the present setup.

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

Moscarini, G. and Smith, L. (2001). The optimal level of experimentation. *Econometrica* 69(6):1629–1644.

Nyarko, Y. (1994). On the convexity of the value function in bayesian optimal control problems. *Economic Theory* 2(2):303–309.

Romer, P. (1986). Cake eating, chattering, and jumps: Existence results for variational problems. *Econometric* 54(4):897–908.

Weber, R. (1992). On the gittins index for multiarmed bandits. *The Annals of Applied Probability* 2(4).

Whittle, P. (1980). Multi-armed bandits and the gittins index. *Journal of the Royal Statistical Society. Series B (Methodological)* 42(2):143–149.

A Proofs

Proof of Theorem 1. It is well known that the value function in Bayesian control problems is convex (see Nyarko, 1994 for the discrete-time case). Since $w(0) = w(1)$, convexity means that w is U-shaped. Therefore, we can find $0 < \underline{\pi} \leq \overline{\pi} < 1$ such that w is non-increasing on $[0, \underline{\pi}]$, constant on $[\underline{\pi}, \overline{\pi}]$, and non-decreasing on $[\overline{\pi}, 1]$. Higher values of π are good news for tool x_0 and bad news for state x_1. Thus, the non-increasing portion of w corresponds to beliefs that recommend $\{x_1\}$, while the non-decreasing region, to beliefs that recommend $\{x_0\}$. $\square$

Lemma A1. *The function $w^0 : [0,1] \to \mathbb{R}$ in* (3) *is continuously differentiable, strictly decreasing below $\underline{\pi}^\beta$, and strictly increasing above $\overline{\pi}^\beta$.*

Proof. Continuous differentiability follows from value matching and smooth pasting. On $\left[0, \underline{\pi}^\beta\right)$,

$$
\begin{aligned}
w^{0\prime}(\pi) &= \frac{\lambda^0 c \psi(\underline{\pi}^\beta)}{\lambda^0 + \rho} \left(\frac{\psi(\pi)}{\psi(\underline{\pi}^\beta)}\right)^{-\frac{\rho}{\lambda^0}} \frac{\lambda^0(1-\pi) + \rho}{\lambda^0(1-\pi)} - \lambda^0 \\
&< \frac{\lambda^0 c \psi(\underline{\pi}^\beta)}{\lambda^0 + \rho} \frac{\lambda^0(1-\underline{\pi}^\beta) + \rho}{\lambda^0(1-\underline{\pi}^\beta)} - \lambda^0 = 0.
\end{aligned}
$$

Finally, on $\left(\overline{\pi}^\beta, 1\right]$,

$$
\begin{aligned}
w^{0\prime}(\pi) &= -\frac{\lambda^0 c}{(\lambda^0 + \rho)\psi(\overline{\pi}^\beta)} \left(\frac{\psi(\pi)}{\psi(\overline{\pi}^\beta)}\right)^{\frac{\rho}{\lambda^0}} \frac{\lambda^0 \pi + \rho}{\lambda^0 \pi} + \lambda^0 \\
&> -\frac{\lambda^0 c}{(\lambda^0 + \rho)\psi(\overline{\pi}^\beta)} \frac{\lambda^0 \overline{\pi}^\beta + \rho}{\lambda^0 \overline{\pi}^\beta} + \lambda^0 = 0.
\end{aligned}
$$

This concludes the proof. $\square$

Proof of Theorem 2. Consider the case $\rho(2c-\lambda^0) > \lambda^0(\lambda^0 - c)$. We want to verify that w^0 in (3) solves the Bellman equation. To this end, define:

$$R^0_{w^0}(\pi) := \lambda^0\pi - c + \frac{\lambda^0\pi\left[\lambda^0 - c - w^0(\pi)\right] - \lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho};$$
$$R^1_{w^0}(\pi) := \lambda^0(1-\pi) - c + \frac{\lambda^0(1-\pi)\left[\lambda^0 - c - w^0(\pi)\right] + \lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho}.$$

We must check the following conditions:

1. On $\left[0, \underline{\pi}^\beta\right)$, $R^1_{w^0}(\pi) > \max\left\{R^0_{w^0}(\pi), 0\right\}$;
2. On $\left(\underline{\pi}^\beta, \overline{\pi}^\beta\right)$, $0 > \max\left\{R^0_{w^0}(\pi), R^1_{w^0}(\pi)\right\}$;
3. Finally, on $\left(\overline{\pi}^\beta, 1\right]$, $R^0_{w^0}(\pi) > \max\left\{R^1_{w^0}(\pi), 0\right\}$.

Start with $\pi \in (0, \underline{\pi}^\beta)$. Using (1), we can write $R^1_{w^0}(\pi) - R^0_{w^0}(\pi)$ as:

$$\begin{aligned}
R^1_{w^0}(\pi) - R^0_{w^0}(\pi) &= \lambda^0(1-2\pi) + \frac{\lambda^0(1-2\pi)\left[\lambda^0 - c - w^0(\pi)\right] + 2\lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho} \\
&= \lambda^0 - \frac{(\lambda^0 + 2\rho)\left[\lambda^0 - c - w^0(\pi)\right]}{\rho} \\
&> \lambda^0 - \frac{(\lambda^0 + 2\rho)\left[\lambda^0 - c - w^0(\underline{\pi}^\beta)\right]}{\rho} \\
&= \lambda^0 - \frac{(\lambda^0 + 2\rho)}{\rho}(\lambda^0 - c) > 0,
\end{aligned}$$

where the first strict inequality follows from the fact that w^0 is strictly decreasing on $[0, \underline{\pi}^\beta)$. Similarly,

$$\begin{aligned}
R^1_{w^0}(\pi) &= \lambda^0(1-\pi) - c + \frac{\lambda^0(1-\pi)\left[\lambda^0 - c - w^0(\pi)\right] + \lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho} \\
&= \lambda^0(1-\pi) - c - \left[\lambda^0 - c - w^0(\pi)\right] + \lambda^0\pi \\
&= w^0(\pi) > w^0(\underline{\pi}^\beta) = 0.
\end{aligned}$$

Next, consider $\pi \in \left(\underline{\pi}^\beta, \overline{\pi}^\beta\right)$. In this region, $w^0(\pi) = w^{0\prime}(\pi) = 0$. Now,

$$R^0_{w^0}(\pi) = \lambda^0\pi\left(\frac{\rho + \lambda^0 - c}{\rho}\right) - c < \lambda^0\overline{\pi}^\beta\left(\frac{\rho + \lambda^0 - c}{\rho}\right) - c = 0;$$
$$R^1_{w^0}(\pi) = \lambda^0(1-\pi)\left(\frac{\rho + \lambda^0 - c}{\rho}\right) - c < \lambda^0\left(1 - \underline{\pi}^\beta\right)\left(\frac{\rho + \lambda^0 - c}{\rho}\right) - c = 0.$$

Finally, for $\pi \in \left(\overline{\pi}^{\beta}, 1\right]$, using (2), we have:

$$\begin{aligned}
R^0_{w^0}(\pi) - R^1_{w^0}(\pi) &= \lambda^0(2\pi - 1) + \frac{\lambda^0(2\pi-1)\left[\lambda^0 - c - w^0(\pi)\right] - 2\lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho} \\
&= \lambda^0 - \frac{\lambda^0 + 2\rho}{\rho}\left[\lambda^0 - c - w^0(\pi)\right] \\
&> \lambda^0 - \frac{\lambda^0 + 2\rho}{\rho}\left[\lambda^0 - c - w^0\left(\overline{\pi}^{\beta}\right)\right] \\
&= \lambda^0 - \frac{(\lambda^0 + 2\rho)}{\rho}(\lambda^0 - c) > 0,
\end{aligned}$$

where the (first) strict inequality follows because w^0 is strictly increasing on this region; and:

$$R^0_{w^0}(\pi) = \lambda^0\pi - c + \frac{-\rho(\lambda^0 - c - w^0(\pi)) + \lambda^0\rho(1-\pi)}{\rho} = w^0(\pi) > w^0\left(\overline{\pi}^{\beta}\right) = 0.$$

The proof that w^0 in (4) solves the Bellman equation, corresponding to the case $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$, is analogous; the details are omitted. In this case, the DM cannot profit from giving up, as $w^0(\pi) \geq w^0(1/2) \geq 0$ if $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$. □

Proof of Theorem 3. Consider the case $\lambda^0(\lambda^0 - c) > \rho(2c - \lambda^0)$. To verify w^0 in (7) solves the Bellman equation, define:

$$R^2_{w^0}(\pi) := \lambda^0 - 2c + \frac{\lambda^0\left[\lambda^0 - c - w^0(\pi)\right]}{\rho}.$$

By assumption, $R^2_{w^0}(\pi) > 0$ for all $\pi \in [0, 1]$. We must check that:

1. On $(0, \underline{\pi}^*)$, $R^1_{w^0}(\pi) - R^0_{w^0}(\pi) > 0$ and $R^1_{w^0}(\pi) - R^2_{w^0}(\pi) > 0$.
2. On $(\underline{\pi}^*, \overline{\pi}^*)$, $R^2_{w^0}(\pi) - R^0_{w^0}(\pi) > 0$ and $R^2_{w^0}(\pi) - R^1_{w^0}(\pi) > 0$.
3. Finally, on $(\overline{\pi}^*, 1)$, $R^0_{w^0}(\pi) - R^1_{w^0}(\pi) > 0$ and $R^0_{w^0}(\pi) - R^2_{w^0}(\pi) > 0$.

Start with $\pi \in (0, \underline{\pi}^*)$. The first inequality in 1 is established in the same way as in the proof of Theorem 2. As for the second,

$$R^1_{w^0}(\pi) - R^2_{w^0}(\pi) = c - \frac{(\lambda^0 + \rho)\left[\lambda^0 - c - w^0(\pi)\right]}{\rho} > c - \frac{(\lambda^0 + \rho)\left[\lambda^0 - c - w^0(\underline{\pi}^*)\right]}{\rho} = 0.$$

Next, take $\pi \in (\underline{\pi}^*, \overline{\pi}^*)$. Now,

$$\begin{aligned}
R^2_{w^0}(\pi) - R^0_{w^0}(\pi) &= \lambda^0(1-\pi)\frac{\lambda^0 + \rho + c}{\lambda^0 + \rho} - c > \lambda^0(1 - \overline{\pi}^*)\frac{\lambda^0 + \rho + c}{\lambda^0 + \rho} - c = 0; \\
R^2_{w^0}(\pi) - R^1_{w^0}(\pi) &= \lambda^0\pi\frac{\lambda^0 + \rho + c}{\lambda^0 + \rho} - c > \lambda^0\overline{\pi}^*\frac{\lambda^0 + \rho + c}{\lambda^0 + \rho} - c = 0.
\end{aligned}$$

As for the second inequality in 3,

$$
\begin{aligned}
R^0_{w^0}(\pi) - R^2_{w^0}(\pi) &= -[\lambda^0(1-\pi) - c] - \frac{\lambda^0(1-\pi)\left[\lambda^0 - c - w^0(\pi)\right] + \lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho} \\
&= c - \frac{(\lambda^0+\rho)\left[\lambda^0 - c - w^0(\pi)\right]}{\rho} \\
&> c - \frac{(\lambda^0+\rho)\left[\lambda^0 - c - w^0(\overline{\pi}^*)\right]}{\rho} = 0.
\end{aligned}
$$

In the case where $\lambda^0(\lambda^0 - c) \leq \rho(2c - \lambda^0)$, it remains to check that the DM cannot profit by choosing the (now available) complete toolkit. This holds because $(2c-\lambda^0)\rho \geq \lambda^0(\lambda^0 - c)$ implies that $\lambda^0 - c - \frac{\rho c}{\lambda^0+\rho} \leq 0$. $\square$

Lemma A2. *The function $w^0 : [0,1] \to \mathbb{R}$ given in* (11) *is continuously differentiable and satisfies $w^{0\prime}(\pi) < \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}$ on $[0, \underline{\pi}^*)$ and $w^{0\prime}(\pi) > \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}$ on $(\overline{\pi}^*, 1]$.*

Proof. Continuous differentiability follows from value matching and smooth pasting. On $[0, \underline{\pi}^*)$,

$$
\begin{aligned}
w^{0\prime}(\pi) &= \frac{\lambda^0 c_0 \psi(\underline{\pi}^*)}{\lambda^0+\rho}\left(\frac{\psi(\pi)}{\psi(\underline{\pi}^*)}\right)^{-\frac{\rho}{\lambda^0}}\left(1 + \frac{\rho}{\lambda^0(1-\pi)}\right) - \lambda^0 \\
&< \frac{\lambda^0 c_0 \psi(\underline{\pi}^*)}{\lambda^0+\rho}\left(1 + \frac{\rho}{\lambda^0(1-\overline{\pi}^*)}\right) - \lambda^0 = \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho};
\end{aligned}
$$

on $(\overline{\pi}^*, 1]$,

$$
\begin{aligned}
w^{0\prime}(\pi) &= -\frac{\lambda^0 c_1}{(\lambda^0+\rho)\psi(\overline{\pi}^*)}\left(\frac{\psi(\pi)}{\psi(\overline{\pi}^*)}\right)^{\frac{\rho}{\lambda^0}}\left(1 + \frac{\rho}{\lambda^0\pi}\right) + \lambda^0 \\
&> -\frac{\lambda^0 c_1}{(\lambda^0+\rho)\psi(\overline{\pi}^*)}\left(1 + \frac{\rho}{\lambda^0\overline{\pi}^*}\right) + \lambda^0 = \frac{\lambda^0(c_1-c_0)}{\lambda^0+\rho}.
\end{aligned}
$$

This concludes the proof. $\square$

Proof of Theorem 4. Start with the portion of theorem regarding the counterpart of Theorem 3. Some special care needs to be taken compared to the argument behind Theorem 3, as the value function is non-monotonic on $(\overline{\pi}^*, 1]$. Define:

$$
\begin{aligned}
R^0_{w^0}(\pi) &:= \lambda^0\pi - c_0 + \frac{\lambda^0\pi\left[\lambda^0 - c_0 - w^0(\pi)\right] - \lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho}; \\
R^1_{w^0}(\pi) &:= \lambda^0(1-\pi) - c_1 + \frac{\lambda^0(1-\pi)\left[\lambda^0 - c_1 - w^0(\pi)\right] + \lambda^0\pi(1-\pi)w^{0\prime}(\pi)}{\rho}; \\
R^2_{w^0}(\pi) &:= \lambda^0 - c_0 - c_1 + \frac{\lambda^0\left[\lambda^0 - c_1 - \pi(c_0-c_1) - w^0(\pi)\right]}{\rho}.
\end{aligned}
$$

We must check the following conditions:

1. On $[0,\underline{\pi}^*)$, $R^1_{w^0}(\pi) - R^0_{w^0}(\pi) > 0$ and $R^2_{w^0}(\pi) - R^1_{w^0}(\pi) < 0$.
2. On $(\overline{\pi}^*,1]$, $R^1_{w^0}(\pi) - R^0_{w^0}(\pi) < 0$ and $R^2_{w^0}(\pi) - R^0_{w^0}(\pi) < 0$.
3. Finally, on $(\underline{\pi}^*,\overline{\pi}^*)$, $R^2_{w^0}(\pi) - R^0_{w^0}(\pi) > 0$ and $R^2_{w^0}(\pi) - R^1_{w^0}(\pi) > 0$.

Start with $\pi \in (0,\underline{\pi}^*)$; the counterpart of (1) is

$$-\lambda^0\pi(1-\pi)w^{0\prime}(\pi) + [\lambda^0(1-\pi)+\rho]w^0(\pi) = \lambda^0(1-\pi)(\rho+\lambda^0-c_1) - \rho c_1.$$

Proceeding as in the symmetric case, we can write:

$$R^1_{w^0}(\pi) - R^0_{w^0}(\pi) = \lambda^0 + c_0 + c_1 - \frac{(\lambda^0+2\rho)\left[\lambda^0 - w^0(\pi)\right]}{\rho} + \frac{\lambda^0 c_1 + \lambda^0\pi(c_0-c_1)}{\rho}.$$

By Lemma A2, as $c_0 > c_1$, the right-hand side of the equation above is strictly decreasing in π (despite the last term being strictly increasing in π). In this case,

$$R^1_{w^0}(\pi) - R^0_{w^0}(\pi) > \lambda^0 + c_0 + c_1 - \frac{(\lambda^0+2\rho)\left[\lambda^0 - w^0(\underline{\pi}^*)\right]}{\rho} + \frac{\lambda^0 c_1 + \lambda^0\underline{\pi}^*(c_0-c_1)}{\rho} = w^0(\underline{\pi}^*);$$

by assumption, $w^0(\underline{\pi}^*) > 0$. Next, take:

$$R^1_{w^0}(\pi) - R^2_{w^0}(\pi) = c_0 - \frac{(\lambda^0+\rho)(\lambda^0 - c_1 - w^0(\pi)) - \lambda^0\pi(c_0-c_1)}{\rho}.$$

From Lemma A2, it follows that the right-hand side is strictly decreasing in π. Thus,

$$R^1_{w^0}(\pi) - R^2_{w^0}(\pi) > c_0 - \frac{(\lambda^0+\rho)(\lambda^0 - c_1 - w^0(\underline{\pi})) - \lambda^0\underline{\pi}(c_0-c_1)}{\rho} = 0.$$

Next, consider $\pi \in (\underline{\pi}^*,\overline{\pi}^*)$. Here, we have:

$$\begin{aligned}
R^2_{w^0}(\pi) - R^0_{w^0}(\pi) &= \lambda^0(1-\pi)\frac{\lambda^0+\rho+c_0}{\lambda^0+\rho} - c_1 \\
&> \lambda^0(1-\overline{\pi}^*)\frac{\lambda^0+\rho+c_0}{\lambda^0+\rho} - c_1 = 0; \\
R^2_{w^0}(\pi) - R^1_{w^0}(\pi) &= \lambda^0\frac{\lambda^0+\rho+c_1}{\lambda^0+\rho}\pi - c_0 \\
&> \lambda^0\frac{\lambda^0+\rho+c_1}{\lambda^0+\rho}\underline{\pi}^* - c_0 = 0.
\end{aligned}$$

Finally, take $\pi \in (\overline{\pi}^*,1]$; (2) is now $\lambda^0\pi(1-\pi)w'(\pi) + (\rho+\lambda^0\pi)w(\pi) = \lambda^0\pi(\rho+\lambda^0-c_0) - \rho c_0$. We have:

$$R^0_{w^0}(\pi) - R^1_{w^0}(\pi) = \lambda^0 + c_0 - c_1 - \frac{\lambda^0+2\rho}{\rho}[\lambda^0 - c_1 - w^0(\pi)] + \frac{\lambda^0(c_0-c_1)}{\rho}\pi.$$

Unlike in the symmetric case, this expression is not monotonic over the range $(\overline{\pi}^*, 1]$. However, if $\underline{w}^0$ denotes the minimum of w^0, we have:

$$R^0_{w^0}(\pi) - R^1_{w^0}(\pi) > \lambda^0 + c_0 - c_1 - \frac{\lambda^0 + 2\rho}{\rho}[\lambda^0 - c_1 - \underline{w}^0] + \frac{\lambda^0(c_0 - c_1)}{\rho}\overline{\pi}^*;$$

by assumption, this last expression is non-negative. As for $R^0_{w^0}(\pi) - R^1_{w^0}(\pi)$, we have:

$$R^0_{w^0}(\pi) - R^2_{w^0}(\pi) = c_0 - \frac{\rho + \lambda^0}{\lambda^0}(\lambda^0 - c_1 - w^0(\pi)) + \frac{\lambda^0}{\rho}(c_0 - c_1)\pi.$$

By Lemma A2, this expression is indeed strictly increasing in π; hence,

$$R^0_{w^0}(\pi) - R^2_{w^0}(\pi) > c_0 - \frac{\rho + \lambda^0}{\lambda^0}(\lambda^0 - c_1 - w^0(\overline{\pi}^*)) + \frac{\lambda^0}{\rho}(c_0 - c_1)\overline{\pi}^* = 0.$$

For the portion of the proof corresponding to the statement about Theorem 2, it remains to check that $R^1_{w^0}(\pi) > 0$ on $[0, \underline{\pi}^\beta)$, that $R^0_{w^0}(\pi) > 0$ on $(\underline{\pi}^\beta, \overline{\pi}^\beta)$, and that $R^0_{w^0}(\pi), R^1_{w^0}(\pi) < 0$ on $(\overline{\pi}^\beta, 1]$. The argument here is completely analogous to the symmetric case, since the regions on which the value function is strictly decreasing, constant, and strictly increasing, correspond to the ranges over which the strategy recommends choosing $\{x_1\}$, $\emptyset$, and $\{x_0\}$, respectively. The details are omitted. $\square$

Lemma A3. *There exists a unique* $\underline{\pi}^r \in \left(0, \min\left\{\overline{\pi}^\beta, \underline{\pi}^\beta\right\}\right)$ *such that:*

$$\frac{\lambda^0 c}{\lambda^0 + \rho}\underline{\pi}^r \left(\frac{\psi(\underline{\pi}^r)}{\psi\left(\max\left\{\overline{\pi}^\beta, \underline{\pi}^\beta\right\}\right)}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1 - \underline{\pi}^r) - c = \lambda^0 - c - \frac{\rho c}{\lambda^0 + \rho};$$

similarly, there exists a unique $\overline{\pi}^r \in \left(\max\left\{\overline{\pi}^\beta, \underline{\pi}^\beta\right\}\right), 1$ *such that:*

$$\frac{\lambda^0 c}{\lambda^0 + \rho}(1 - \overline{\pi}^r) \left(\frac{\psi(\overline{\pi}^r)}{\psi\left(\min\left\{\overline{\pi}^\beta, \underline{\pi}^\beta\right\}\right)}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\overline{\pi}^r - c = \lambda^0 - c - \frac{\rho c}{\lambda^0 + \rho}.$$

Proof. Consider the case $\rho(2c - \lambda^0) \le \lambda^0(\lambda^0 - c)$; the other case is handled analogously. Define the following function $h : [0, 1] \to \mathbb{R}$, given by:

$$h(x) := \frac{\lambda^0 c}{\lambda^0 + \rho} x \left(\frac{\psi(x)}{\psi\left(\underline{\pi}^\beta\right)}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1 - x) - c - (\lambda^0 - c) + \frac{\rho c}{\lambda^0 + \rho}.$$

By Lemma A1, h is differentiable and strictly decreasing on $\left[0, \underline{\pi}^\beta\right)$. Moreover, this function satisfies:

$$h(0) = \frac{\rho c}{\lambda^0 + \rho} > 0;$$
$$h\left(\overline{\pi}^\beta\right) < \frac{\lambda^0 c}{\lambda^0 + \rho}\overline{\pi}^\beta + \lambda^0 \underline{\pi}^\beta - c - (\lambda^0 - c) + \frac{\rho c}{\lambda^0 + \rho} = 0.$$

Thus, there exists a unique $x^* \in \left(0, \overline{\pi}^\beta\right)$ such that $h(x^*) = 0$. A similar argument as above establishes that there exists a unique $x^{**} \in \left(\underline{\pi}^\beta, 1\right)$ such that $g(x^{**}) = 0$, where $g : [0,1] \to \mathbb{R}$ is given by:

$$g(x) := \frac{\lambda^0 c}{\lambda^0 + \rho}(1-x)\left(\frac{\psi(x)}{\psi\left(\overline{\pi}^\beta\right)}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0 x - c - (\lambda^0 - c) + \frac{\rho c}{\lambda^0 + \rho}.$$

Set $\underline{\pi}^r = x^*$ and $\overline{\pi}^r = x^{**}$. □

Proof of Theorem 5. There is nothing to show if the feasible set is the empty set. The portions of the theorem corresponding to singletons being the feasible sets follow as in the proof of Theorem 2. (The only difference is that, here, we do not need to worry about having $\underline{\pi}^\beta < \overline{\pi}^\beta$; the two cutoffs apply to different states.) As for the last two cases, it suffices to compare the value functions in (7) and (3) to the value corresponding to the complete toolkit. Start with the case $\rho(2c - \lambda^0) \leq \lambda^0(\lambda^0 - c)$. We have $\underline{\pi}^* > \overline{\pi}^\beta$; thus, on $[0, \overline{\pi}^\beta]$, $w^0(\pi) - (\lambda^0 - c) + \frac{\rho c}{\lambda^0+\rho} = h(\pi)$, where h is as in the proof of Lemma A3. Thus, for all $\pi < \underline{\pi}^r$, $w^0(\pi) > \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho}$. Similarly, on $[\underline{\pi}^\beta, 1]$, we have $w^0(\pi) - (\lambda^0 - c) + \frac{\rho c}{\lambda^0+\rho} = g(\pi)$, and (the proof of) Lemma A3 establishes that $w^0(\pi) > \lambda^0 - c - \frac{\rho c}{\lambda^0+\rho}$ for all $\pi > \overline{\pi}^r$. Finally, if $\lambda^0(\lambda^0 - c) < \rho(2c - \lambda^0)$, the desired result follows from the fact that $\lambda^0 - c - \frac{\rho c}{\lambda^0+\rho} < 0$. □

Proof of Theorem 6. Consider a cutoff strategy of the same form as in the beginning of Section 4. We have the same differential equation as before. The new (VM) conditions are $w(\underline{\pi}) = \lambda^0 - 2c = w(\overline{\pi})$. These conditions lead to

$$C_1 = C_1(\underline{\pi}) := \frac{\lambda^0 \underline{\pi} - c}{\underline{\pi}}\psi(\underline{\pi})^{\frac{\rho}{\lambda^0}}, \quad C_0 = C_0(\overline{\pi}) := \frac{\lambda^0(1-\overline{\pi}) - c}{(1-\overline{\pi})}\psi(\overline{\pi})^{-\frac{\rho}{\lambda^0}},$$

$\underline{\pi} = \frac{\lambda^0+\rho}{\rho+c}\frac{c}{\lambda^0} \in (0,1)$, and $\overline{\pi} = 1 - \underline{\pi}$. We have $\overline{\pi} > \underline{\pi}$ if and only if $\rho(\lambda^0 - 2c) > \lambda^0 c$. The solution candidate in this case is:

$$w^0(\pi) = \begin{cases} \frac{\lambda^0(\lambda^0-c)}{\lambda^0+\rho}\pi\left(\frac{\psi(\pi)}{\psi(\underline{\pi})}\right)^{-\frac{\rho}{\lambda^0}} + \lambda^0(1-\pi) - c & \pi \in [0, \underline{\pi}); \\ \lambda^0 - 2c & \pi \in [\underline{\pi}, \overline{\pi}]; \\ \frac{\lambda^0(\lambda^0-c)}{\lambda^0+\rho}(1-\pi)\left(\frac{\psi(\pi)}{\psi(\overline{\pi})}\right)^{\frac{\rho}{\lambda^0}} + \lambda^0\pi - c & \pi \in (\overline{\pi}, 1]. \end{cases}$$

The proof that this function solves the Bellman equation is entirely analogous to the corresponding proof in Theorem 3. The argument for the case $\rho(\lambda^0 - 2c) \leq \lambda^0 c$ is analogous to the argument behind Theorem 2; notice that:

$$w^0\left(\frac{1}{2}\right) = \frac{\lambda^0(\rho + \lambda^0 - c) - 2\rho c}{\lambda^0 + 2\rho} \geq \lambda^0 - 2c$$

if and only if $-\lambda^0(\lambda^0 - c) \leq \rho(\lambda^0 - 2c) \leq \lambda^0 c$, and $w^0(1/2) < 0$ if and only if $-\lambda^0(\lambda^0 - c) > \rho(\lambda^0 - 2c)$. $\square$

Proof of Theorem 7. The value function is:

$$w(\pi, \mu) = \begin{cases} w^0(\pi) & \mu < \mu(\pi); \\ \frac{\lambda^1}{\lambda^1+\rho}(1-\mu)\left(\frac{\psi(\mu)}{\psi(\mu(\pi))}\right)^{\frac{\rho}{\lambda^1}} + \lambda^1\mu - c_y & \mu \geq \mu(\pi). \end{cases}$$

Fix $\pi \in [0,1]$. By the same argument as in Lemma A1, $w(\pi, \mu)$ is strictly increasing in μ on $[\mu(\pi), 1]$, and attains the value $w^0(\pi)$ at $\mu = \mu(\pi)$. Thus, this function attains the maximum in the Bellman equation. $\square$