Online Appendix for Efficient Multi-Agent Experimentation and Multi-Choice Bandits: Extensions

This online appendix extends the analysis to the cases where the decision to neglect projects is irreversible (e.g. if projects are scooped; section 1); where joint research "destroys" information because the specific source of success cannot be identified (section 2); and where there is a third, independent but riskier project (section 3). Propositions follow the numbering of the main text; proofs are given at the end (section 4).

1. Nested Choices

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. When choices must be nested, there is an option value to holding on to projects beyond what an unrestricted DM would consider optimal.

Proposition 3 (Nested choices)**.** *The optimal strategy when choices are nested is as follows. There are two unique cutoffs $\underline{\pi}^N \in (0, \min\{1-\underline{\pi}^1, \underline{\pi}^1\})$ and $\overline{\pi}^N \in (\max\{1-\overline{\pi}^1, \underline{\pi}^1\}, 1)$ such that:*

- *If currently working only on project 1, stick to it if $\pi \leq \underline{\pi}^1$ (and give up otherwise);*
- *If currently working only on project 0, stick to it if $\pi \geq \overline{\pi}^1$ (and give up otherwise);*
- *Under costly research, choose as in Proposition 1;*
- *Under beneficial research, work on 1 if $\pi < \underline{\pi}^N$, on 0 if $\pi > \overline{\pi}^N$, and otherwise on both at once.*

If research is costly, a sufficiently impatient DM behaves as her unrestricted counterpart does. However, under beneficial research, the DM has to be more certain about the state to focus on a single project: If $\rho(2c-\overline{\lambda}) \leq \overline{\lambda}(\overline{\lambda}-c)$, then $\underline{\pi}^N < \overline{\pi}^1 \leq \underline{\pi}^2 \leq \overline{\pi}^2 \leq \underline{\pi}^1 < \overline{\pi}^N$. Intuitively, by focusing on a single project, the DM is giving up the option value of being able to switch to the other project at a later point in time — after having gathered more information.

In the context of multi-agent experimentation, restricting choices to be nested means that the action of assigning a player to their safe arm is irreversible. Once the social planner decides a player should take her safe option, she cannot reassign them to experimenting in the future.

2. Imperfect Monitoring of Successes

In recruiting, it may be the case that individual contributions to team output cannot be readily assessed. In labs, when testing two different treatments, it may be difficult to control for some chemical interactions that may obscure the findings. If the DM, when doing simultaneous research, can only observe the occurrence of arrivals but not their "precedence," then simultaneous research is as uninformative as is not doing research. To gather information, the DM must give individual projects a chance *on their own*; "experimentation" now entails focusing on a project at a time.

The relevant cases of costly and beneficial research are now as follows.

Costly research 2. $\rho(\overline{\lambda} - 2c) > \overline{\lambda}c$.
Beneficial research 2. $\rho(\overline{\lambda} - 2c) \leq \overline{\lambda}c$.

Under costly research 2, the cost of undertaking projects is low ($c < \overline{\lambda}/2$), and the DM is impatient ($\rho > \overline{\lambda}c/(\overline{\lambda} - 2c)$). Thus, the temptation to forgo information and exploit the projects simultaneously is high. Under beneficial research 2, either the cost of undertaking projects is high, or the DM is sufficiently patient. Thus, working on a single project is an appealing proposition. Notice that beneficial research 2 is consistent with both costly and beneficial research as in the main text.

Proposition 4 (Imperfect monitoring)**.** *The optimal strategy under imperfect monitoring of successes is as follows. Under costly research 2, we have: Work on 1 if* $\pi < \underline{\pi}^{IM}$, *on 0 if* $\pi > 1 - \underline{\pi}^{IM}$, *and on both at once otherwise, where* $\underline{\pi}^{IM} := \frac{\overline{\lambda}+\rho}{\rho+c}\frac{c}{\overline{\lambda}}$. *Under beneficial research 2, the optimal strategy equals that of costly research in Proposition 1 if* $-\overline{\lambda}(\overline{\lambda} - c) > \rho(\overline{\lambda} - 2c)$; *and if* $-\overline{\lambda}(\overline{\lambda} - c) \leq \rho(\overline{\lambda} - 2c)$, *it recommends working on 1 if* $\pi < 1/2$, *on 0 if* $\pi > 1/2$, *and alternating evenly between the two if* $\pi = \frac{1}{2}$.

If research is beneficial, an impatient DM engages in simultaneous research and enjoys a constant expected payoff when she is sufficiently unsure about the state. Conversely, when research is expensive or the DM is sufficiently patient, the information that only singletons can provide is valuable. The social planner in Klein and Rady (2011) (KR) would face this problem if both players' risky arms belonged to the same bandit, so that only the arrival of a success, and not the source, can be identified. In this case, more information is generated by having the players experiment one at a time, and a patient social planner would have them do just that.

3. A Third, Unrelated Risky Project

Imagine that the DM has a third project on which she can work, labelled project 2. If this third project is productive, its arrival rate is higher than that of the other two. But this new project may be unproductive, while one of the original two projects is certainly productive. To keep the problem simple, I assume that this third project is "incompatible" with the other two in the sense that it requires the full attention of the DM while she is working on it, and that it must be forsaken once ignored or abandoned. Thus, the problem is to determine for how long to experiment on the riskier project, if at all, before switching to one or both of the original two.

Undertaking the new project also involves a flow cost of c; but if the new project is fruitful, it produces successes at rate $\overline{\lambda}_2 > \overline{\lambda}$. Thus, the DM knows that, if this new project proves successful, it is more appealing than any of the others; otherwise, she is better off with the original ones. Let $\mu \in [0,1]$ denote the assessment of the DM that the new project is fruitful; $\mu^0 \in [0,1]$ denotes the corresponding prior. Representing separate projects, I assume that the new project is independent of the original two. Under irreversibility, once the DM switches away from 2, she is back to our basic problem.

Proposition 5 (Third project)**.** *The optimal strategy dictates starting on 2 provided that $\mu^0 \geq \underline{\mu}(\pi^0)$, and sticking to it while the posterior (π, μ) satisfies $\mu \geq \underline{\mu}(\pi)$, where $\underline{\mu}(\pi) := \frac{\rho}{\overline{\lambda}_2} \frac{w(\pi)+c}{\overline{\lambda}_2+\rho-w(\pi)-c}$ and $w(\pi)$ is the value function from the strategy in Proposition 1; and if $\mu < \underline{\mu}(\pi)$, switching to what Proposition 1 dictates.*

If the DM is sufficiently sure about which of the original two projects will work out, she has to be sufficiently confident about project 2 to start working on it. Otherwise, she sticks to project 2 for a wider range of beliefs: If she switches to the original projects being uncertain about them, she will work on both at once and bear a higher total research cost, or give up altogether.

Going back to KR, the third project can be seen as a third player with a riskier arm of her own. If this new player's risky arm is more appealing, the social planner holds out on the original players in favor of the riskier player experimenting first.

4. Proofs

4.1. Nested Choices

Let $w^\alpha : [0,1] \times 2^{\{0,1\}} \to \mathbb{R}$ represent the value function augmented to include the set of feasible projects as a second state. Clearly, $w^\alpha(\pi, \emptyset) = 0$. The Bellman equation for $w^\alpha(\pi, \{1\})$ is:

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

the same (VM) and (SP) conditions relating the choice of 1 and the empty set apply. Thus, we have:

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

The same argument applies to $w^\alpha(\pi, \{0\})$, leading to:

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

Finally, for $w^\alpha(\pi, \{0,1\})$, we have:

$$\begin{aligned} &w^\alpha(\pi, \{0,1\}) = \\ &\max\left\{ w^\alpha(\pi, \{0\}), w^\alpha(\pi, \{1\}), \left(\overline{\lambda} - 2c + \frac{\overline{\lambda}(\overline{\lambda} - c - w^\alpha(\pi, \{0,1\}))}{\rho} - w^\alpha(\pi, \{0,1\})\right) \mathrm{d}t \right. \\ &\left. + w^\alpha(\pi, \{0,1\}) \right\}. \end{aligned}$$

Assume that $\rho(2c - \overline{\lambda}) \leq \overline{\lambda}(\overline{\lambda} - c)$; hence, $\underline{\pi}^1 \geq 1/2 \geq \overline{\pi}^1$. (The case $\underline{\pi}^1 < \overline{\pi}^1$ is handled similarly.) On $[0, \overline{\pi}^1)$,

$$\begin{aligned} &w^\alpha(\pi, \{0,1\}) = \max\{w^\alpha(\pi, \{1\}), \\ &\left(\overline{\lambda} - 2c + \frac{\overline{\lambda}(\overline{\lambda} - c - w^\alpha(\pi, \{0,1\}))}{\rho} - w^\alpha(\pi, \{0,1\})\right) \mathrm{d}t + w^\alpha(\pi, \{0,1\}) \right\} \end{aligned}$$

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

$$\begin{aligned} &w^\alpha(\pi, \{0,1\}) = \\ &\max\left\{ w^\alpha(\pi, \{0\}), \left(\overline{\lambda} - 2c + \frac{\overline{\lambda}(\overline{\lambda} - c - w^\alpha(\pi, \{0,1\}))}{\rho} - w^\alpha(\pi, \{0,1\})\right) \mathrm{d}t \right. \\ &\left. + w^\alpha(\pi, \{0,1\}) \right\}, \end{aligned}$$

and we seek an analogous cutoff $\overline{\pi} \in (\underline{\pi}^1, 1)$ for $w^\alpha(\pi, \{0\})$.

Lemma 1. *There exists a unique* $\underline{\pi}^N \in (0, \min\{\overline{\pi}^1, \underline{\pi}^1\})$ *such that:*

$$\frac{\overline{\lambda} c}{\overline{\lambda}+\rho}\underline{\pi}^N \left(\frac{\Omega(\underline{\pi}^N)}{\Omega(\max\{\overline{\pi}^1, \underline{\pi}^1\})}\right)^{-\frac{\rho}{\overline{\lambda}}} + \overline{\lambda}(1 - \underline{\pi}^N) - c = \overline{\lambda} - c - \frac{\rho c}{\overline{\lambda}+\rho};$$

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

$$\frac{\overline{\lambda} c}{\overline{\lambda}+\rho}(1 - \overline{\pi}^N) \left(\frac{\Omega(\overline{\pi}^N)}{\Omega(\min\{\overline{\pi}^1, \underline{\pi}^1\})}\right)^{\frac{\rho}{\overline{\lambda}}} + \overline{\lambda}\overline{\pi}^N - c = \overline{\lambda} - c - \frac{\rho c}{\overline{\lambda}+\rho}.$$

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

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

This function is differentiable and strictly decreasing on $[0, \underline{\pi}^1)$. Moreover, it satisfies:

$$h(0) = \frac{\rho c}{\overline{\lambda} + \rho} > 0;$$

$$h\left(\overline{\pi}^1\right) < \frac{\overline{\lambda} c}{\overline{\lambda} + \rho}\overline{\pi}^1 + \overline{\lambda}\underline{\pi}^1 - c - (\overline{\lambda} - c) + \frac{\rho c}{\overline{\lambda} + \rho} = 0.$$

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

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

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

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

4.2. Imperfect Monitoring of Successes

Proof of Proposition 4. The (SP) and (VM) conditions for strategies recommending simultaneous research for mid-range beliefs lead to $\underline{\pi} = \frac{\overline{\lambda} + \rho}{\rho + c}\frac{c}{\overline{\lambda}} \in (0, 1)$ and $\overline{\pi} = 1 - \underline{\pi}$. We have $\overline{\pi} > \underline{\pi}$ if and only if $\rho(\overline{\lambda} - 2c) > \overline{\lambda} c$. The solution candidate in this case is:

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

The proof that this function solves the Bellman equation is entirely analogous to the corresponding proof in Proposition 1. So is the proof that w solves the Bellman equation in the case $\rho(\overline{\lambda} - 2c) \leq \overline{\lambda} c$; notice that:

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

if and only if $-\overline{\lambda}(\overline{\lambda} - c) \leq \rho(\overline{\lambda} - 2c) \leq \overline{\lambda} c$, and $w(1/2) < 0$ if and only if $-\overline{\lambda}(\overline{\lambda} - c) > \rho(\overline{\lambda} - 2c)$. □

4.3. A Third, Unrelated Risky Project

While the DM has not switched away from 2, the Bellman equation for the new value function $w_2 : [0,1]^2 \to \mathbb{R}$ is:

$$w_2(\pi,\mu) \\ = \max\left\{ w(\pi), \overline{\lambda}_2\mu - c + \frac{\overline{\lambda}_2\mu\left(w_2\left(\pi,1\right) - w_2(\pi,\mu)\right) - \overline{\lambda}_2\mu(1-\mu)\frac{\partial w_2(\pi,\mu)}{\partial\mu}}{\rho}\right\}.$$

Consider a strategy such that, for each $\pi \in [0,1]$, there is some $\underline{\mu}(\pi) \in [0,1]$ such that the DM starts by working on 2 if $\mu \geq \underline{\mu}(\pi)$, and follows the optimal strategy in Proposition 1 otherwise. The (VM) and (SP) conditions are:

Condition (VM)**.** $w_2\left(\pi, \underline{\mu}(\pi)\right) = w(\pi)$ for all $\pi \in [0,1]$.

Condition (SP)**.** $\frac{\partial w_2\left(\pi,\underline{\mu}(\pi)\right)}{\partial\mu} = 0$ for all $\pi \in [0,1]$.

Proof of Proposition 5. On the region of the state space where the DM experiments with 2, we have:

$$w_2(\pi,\mu) = \overline{\lambda}_2\mu - c + \frac{\overline{\lambda}_2\mu\left(w_2\left(\pi,1\right) - w_2(\pi,\mu)\right) - \overline{\lambda}_2\mu(1-\mu)\frac{\partial w_2(\pi,\mu)}{\partial\mu}}{\rho}.$$

By assumption, $w_2\left(\pi,1\right) = \overline{\lambda}_2 - c$. Thus, $w_2(\pi,\mu) = C(\pi)(1-\mu)\Omega(\mu)^{\frac{\rho}{\overline{\lambda}_2}} + \overline{\lambda}_2\mu - c$, where $C(\cdot)$ is some continuously differentiable function. From the (VM) and (SP) conditions, we find $\underline{\mu}(\pi) = \frac{\rho}{\overline{\lambda}_2}\frac{w(\pi)+c}{\overline{\lambda}_2+\rho-w(\pi)-c}$, and obtain the value function:

$$w_2(\pi,\mu) = \begin{cases} w(\pi) & \mu < \underline{\mu}(\pi); \\ \frac{\overline{\lambda}_2}{\overline{\lambda}_2+\rho}(1-\mu)\left(\frac{\Omega(\mu)}{\Omega(\underline{\mu}(\pi))}\right)^{\frac{\rho}{\overline{\lambda}_2}} + \overline{\lambda}_2\mu - c & \mu \geq \underline{\mu}(\pi). \end{cases}$$

Fix $\pi \in [0,1]$. By standard arguments, $w_2(\pi,\mu)$ is strictly increasing in μ on $[\underline{\mu}(\pi),1]$, and attains the value $w(\pi)$ at $\mu = \underline{\mu}(\pi)$. Thus, this function attains the maximum in the Bellman equation. $\square$

References

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