Online Appendix for Efficient Multi-Agent Experimentation and Multi-Choice Bandits: Further Details and Proofs

This online appendix presents some additional details and proofs. The basic Bellman equation corresponding to section 2 of the article is:

$$\begin{aligned} w(\pi) = \max\Bigg\{ & 0, \overline{\lambda}\pi - c + \frac{\overline{\lambda}\pi\left(w\left(1\right) - w(\pi)\right) - \overline{\lambda}\pi(1-\pi)w'(\pi)}{\rho}, \\ & \overline{\lambda}(1-\pi) - c + \frac{\overline{\lambda}(1-\pi)\left(w\left(0\right) - w(\pi)\right) + \overline{\lambda}\pi(1-\pi)w'(\pi)}{\rho}, \\ & \overline{\lambda} - 2c + \frac{\overline{\lambda}\left(\overline{\lambda} - c - w(\pi)\right)}{\rho} \Bigg\}, \end{aligned} \tag{1}$$

where $w(\pi)$ denotes the value function.

Let $\overline{\pi} \in \left(\frac{1}{2}, 1\right)$ be a cutoff such that, for all beliefs $\pi > \overline{\pi}$, the optimal strategy prescribes experimenting on project 0 alone; and for all $\pi < 1-\overline{\pi}$, the prescribed choice is project 1 alone. Following Klein and Rady (2011), this cutoff is identified by means of the value-matching (VM) and smooth-pasting (SP) conditions:

(**VM**) $w(\overline{\pi}) = w(1-\overline{\pi}) =$

$$\begin{cases} 0 & \text{if no projects are prescribed for } \pi \in [1-\overline{\pi}, \overline{\pi}], \\ \overline{\lambda} - c - \frac{\rho c}{\overline{\lambda}+\rho} & \text{if simultaneous research is prescribed for } \pi \in [1-\overline{\pi}, \overline{\pi}]. \end{cases}$$

(**SP**) $w'(\overline{\pi}) = w'(1-\overline{\pi}) = 0.$

For no research on intermediate beliefs, the VM and SP conditions yield:

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

(with $\overline{\pi}^1 > \frac{1}{2}$ if and only if $\rho(2c-\overline{\lambda}) > \overline{\lambda}(\overline{\lambda}-c)$); for simultaneous research,

$$\overline{\pi}^2 := \frac{\overline{\lambda}(\overline{\lambda}+\rho) - c\rho}{\overline{\lambda}(\overline{\lambda}+\rho+c)}$$

(with $\overline{\pi}^2 > \frac{1}{2}$ if and only if $\rho(2c-\overline{\lambda}) < \overline{\lambda}(\overline{\lambda}-c)$).

For the setting where the bad project is productive (section 3 of the article), we have the Bellman equation:

$$
\begin{aligned}
w(\pi) = \max\Bigg\{ & 0, \Lambda(\pi) - c + \frac{\Lambda(\pi)\left(w\left(j_+(\pi)\right) - w(\pi)\right) - \Delta\lambda\pi(1-\pi)w'(\pi)}{\rho}, \\
& \Lambda(1-\pi) - c + \frac{\Lambda(1-\pi)\left(w\left(j_-(\pi)\right) - w(\pi)\right) + \Delta\lambda\pi(1-\pi)w'(\pi)}{\rho}, \\
& \overline{\lambda} + \underline{\lambda} - 2c + \frac{\Lambda(\pi)w\left(j_+(\pi)\right) + \Lambda(1-\pi)w\left(j_-(\pi)\right) - (\overline{\lambda}+\underline{\lambda})w(\pi)}{\rho} \Bigg\},
\end{aligned} \tag{2}
$$

where, following Keller and Rady (2010), $\Lambda(\pi) := \overline{\lambda}\pi + \underline{\lambda}(1-\pi)$ is the expected arrival rate; $\Delta\lambda := \overline{\lambda} - \underline{\lambda}$ is the difference in arrival rates; and $j_+(\pi) := \frac{\overline{\lambda}\pi}{\Lambda(\pi)}$ and (for the present problem) $j_-(\pi) := 1 - j_+(1-\pi)$ are the jumps in the posterior beliefs upon arrivals.

When the optimal strategy prescribes no research for intermediate beliefs, the solution to the resulting equation is almost identical to Proposition 1 in Keller and Rady (2010). Said strategy is characterized by the cutoff belief:

$$
\overline{\pi}^3 := \frac{(c-\underline{\lambda})\mu}{(\overline{\lambda}-c)(\mu+1) + (c-\underline{\lambda})\mu},
$$

where μ is the positive root of the function $f(x) = \rho + \underline{\lambda} - \Delta\lambda x - \underline{\lambda}\left(\underline{\lambda}/\overline{\lambda}\right)^x$. The difference with Keller and Rady (2010) is that we require that $\overline{\pi}^3 > \frac{1}{2}$.

Proposition (OA1)**.** *We have* $\overline{\pi}^3 > \frac{1}{2}$ *if and only if:* (a) $\overline{\lambda} + \underline{\lambda} < 2c$ *and*

$$
(b) \quad \rho > \Delta\lambda\left(\frac{\overline{\lambda}-c}{2c-\overline{\lambda}-\underline{\lambda}}\right) + \underline{\lambda}\left(\frac{\underline{\lambda}}{\overline{\lambda}}\right)^{\left(\frac{\overline{\lambda}-c}{2c-\overline{\lambda}-\underline{\lambda}}\right)} - \underline{\lambda}.
$$

Proof. Rearranging terms, we find that $\overline{\pi}^3 > \frac{1}{2}$ if and only if $(2c - \overline{\lambda} - \underline{\lambda})\mu > \overline{\lambda} - c$. Since $\overline{\lambda} > c$, we must have that $2c - \overline{\lambda} - \underline{\lambda} > 0$ (condition (a)), and thus $\mu > \frac{\overline{\lambda}-c}{2c-\overline{\lambda}-\underline{\lambda}}$. Since the function $f(x)$ introduced in the last paragraph is strictly decreasing, the last inequality is equivalent to $0 = f(\mu) < f\left(\frac{\overline{\lambda}-c}{2c-\overline{\lambda}-\underline{\lambda}}\right)$ (which leads to (b)). □

Conditions (a) and (b) reduce to the benchmark characterization of costly research provided in the article when $\underline{\lambda} = 0$.

For the next result, we treat $\underline{\lambda}$ as a variable, so write $\mu = \mu(\underline{\lambda})$ and $\overline{\pi}^3 = \overline{\pi}^3(\underline{\lambda})$; notice that $\overline{\pi}^3(0) = \overline{\pi}^1$. All other parameters remain fixed throughout.

Proposition (OA2)**.** $\lim_{\underline{\lambda}\to 0} \overline{\pi}^{3\prime}(\underline{\lambda}) \propto c - \rho$.

Proof. The Implicit Function Theorem implies that $\mu(\underline{\lambda})$ is differentiable, and $\mu'(\underline{\lambda}) > 0$. Thus, $\overline{\pi}^3(\underline{\lambda})$ is also differentiable; we have:

$$
\overline{\pi}^{3\prime}(\underline{\lambda}) = \frac{-(\overline{\lambda}-c)\mu(\underline{\lambda})(1+\mu(\underline{\lambda})) + (\overline{\lambda}-c)(c-\underline{\lambda})\mu'(\underline{\lambda})}{[(\overline{\lambda}-c)(\mu(\underline{\lambda})+1) + (c-\underline{\lambda})\mu(\underline{\lambda})]^2}.
$$

By L'Hopital's rule, we have that $\lim_{\underline{\lambda}\to 0}\mu'(\underline{\lambda}) = \frac{1+\frac{\varrho}{\lambda}}{\overline{\lambda}} = \frac{1+\mu(0)}{\overline{\lambda}}$, and so:

$$\begin{aligned}
\lim_{\underline{\lambda}\to 0}\overline{\pi}^{3\prime}(\underline{\lambda}) &= \frac{-(\overline{\lambda}-c)\mu(0)(1+\mu(0)) + (\overline{\lambda}-c)c\frac{1+\mu(0)}{\overline{\lambda}}}{[\overline{\lambda}-c+\overline{\lambda}\mu(0)]^2} \\
&= \frac{(\overline{\lambda}-c)(1+\mu(0))\frac{c-\rho}{\overline{\lambda}}}{[\overline{\lambda}-c+\overline{\lambda}\mu(0)]^2} \\
&= \frac{(\overline{\lambda}-c)(1+\mu(0))}{\overline{\lambda}[\overline{\lambda}-c+\overline{\lambda}\mu(0)]^2}(c-\rho).
\end{aligned}$$

The result follows. □

The value function when the decision maker (DM) experiments on both projects at once satisfies the following equation:

$$(\rho+\overline{\lambda}+\underline{\lambda})w(\pi) = \rho(\overline{\lambda}+\underline{\lambda}-2c) + \Lambda(\pi)w(j_+(\pi)) + \Lambda(1-\pi)w(j_-(\pi)).$$

Let $\overline{\pi}^4 > \frac{1}{2}$ denote the posterior threshold beyond which the DM focuses on project 0. In this range, the posterior can only jump upward; following Keller and Rady (2010), we have $w(\pi) = \Lambda(\pi) - c + K(1-\pi)\Omega(\pi)^\mu$, where K is a constant of integration and $\Omega(\pi) := \frac{1-\pi}{\pi}$ is the odds ratio. By symmetry, for beliefs below $1-\overline{\pi}^4$, the DM focuses on project 1, the posterior can only jump downward, and we have $w(\pi) = \Lambda(1-\pi) - c + K\pi\Omega(1-\pi)^\mu$.

Posteriors in between $j_+^{-1}(\overline{\pi}^4)$ and $\overline{\pi}^4$ will jump upwards to the region $(\overline{\pi}^4, 1)$; similarly, posteriors in between $1-\overline{\pi}^4$ and $j_-^{-1}(1-\overline{\pi}^4)$ will jump downward to the region $(0, 1-\overline{\pi}^4)$. If the intervals $[1-\overline{\pi}^4, j_-^{-1}(1-\overline{\pi}^4))$ and $(j_+^{-1}(\overline{\pi}^4), \overline{\pi}^4]$ intersect, the value function on the intersection satisfies:

$$\begin{aligned}
(\rho+\overline{\lambda}+\underline{\lambda})w(\pi) = \rho(\overline{\lambda}+\underline{\lambda}-2c)& \\
&+ \Lambda(\pi)\left[\Lambda(j_+(\pi)) - c + K(1-j_+(\pi))\Omega(j_+(\pi))^\mu\right] \\
&+ \Lambda(1-\pi)\left[\Lambda(1-j_-(\pi)) - c + Kj_-(\pi)\Omega(1-j_-(\pi))^\mu\right].
\end{aligned}$$

If the intersection is empty, on the interval $(j_+^{-1}(\overline{\pi}^4), \overline{\pi}^4]$ we have:

$$\begin{aligned}
(\rho+\overline{\lambda}+\underline{\lambda})w(\pi) = \rho(\overline{\lambda}+\underline{\lambda}-2c)& \\
&+ \Lambda(\pi)\left[\Lambda(j_+(\pi)) - c + K(1-j_+(\pi))\Omega(j_+(\pi))^\mu\right] \\
&+ \Lambda(1-\pi)w(j_-(\pi)).
\end{aligned}$$

To determine $w(\pi)$, we must first identify the structure of $w(j_-(\pi))$, which in turn requires identifying the structure of $w(\pi)$ in the neighborhood of $1-\overline{\pi}^4$. Whether said intersection is empty depends on the parameters of the problem.

This construction is carried out until $w(\pi)$ is characterized on the entire interval $[1-\overline{\pi}^4, \overline{\pi}^4]$. The VM and SP conditions pin down the thresholds and constants of integration.

References

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

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