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:

\[
\begin{align*}
    w(\pi) &= \max\left\{ 0, \lambda\pi - c + \frac{\lambda\pi(1-w(1)) - \lambda\pi(1-\pi)w'(\pi)}{\rho}, \right. \\
    &\quad -c + \left. \lambda(1-\pi) - \rho c + \lambda(1-\pi)(w(0) - w(\pi)) + \lambda(1-\pi)w'(\pi) \right\}, \\
    w'(\pi) &= \left( \lambda - 2c + \frac{\lambda(\lambda - c - w(\pi))}{\rho} \right),
\end{align*}
\]

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

Let \(\pi \in \left(\frac{1}{2}, 1\right)\) be a cutoff such that, for all beliefs \(\pi > \pi\), the optimal strategy prescribes experimenting on project 0 alone; and for all \(\pi < 1 - \pi\), 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) \(w(\pi) = w(1 - \pi) = \)

\[
\begin{align*}
    \left\{ \begin{array}{ll}
    0 & \text{if no projects are prescribed for } \pi \in [1 - \pi, \pi], \\
    \lambda - c - \frac{\rho c}{\lambda + \rho} & \text{if simultaneous research is prescribed for } \pi \in [1 - \pi, \pi].
    \end{array} \right.
\end{align*}
\]

(SP) \(w'(\pi) = w'(1 - \pi) = 0\).

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

\[
\bar{\pi}^1 := \frac{cp}{\lambda(\lambda + \rho - c)}
\]

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

\[
\bar{\pi}^2 := \frac{\lambda(\lambda + \rho) - cp}{\lambda(\lambda + \rho + c)}
\]

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

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

\[
\begin{align*}
    w(\pi) = & \max \left\{ 0, \Lambda(\pi) - c + \frac{\Lambda(\pi) (w(j_+(\pi)) - w(\pi)) - \Delta \lambda \pi (1 - \pi) w'(\pi)}{\rho}, \right. \\
    & \left. \Lambda(1 - \pi) - c + \frac{\Lambda(1 - \pi) (w(j_-(\pi)) - w(\pi)) + \Delta \lambda \pi (1 - \pi) w'(\pi)}{\rho}, \right. \\
    & \left. \bar{\lambda} + \Lambda - 2c + \frac{\Lambda(\pi) w(j_+(\pi)) + \Lambda(1 - \pi) w(j_-(\pi)) - (\bar{\lambda} + \Lambda) w(\pi)}{\rho} \right\},
\end{align*}
\]

where, following Keller and Rady (2010), \( \Lambda(\pi) := \bar{\lambda} \pi + \Lambda(1 - \pi) \) is the expected arrival rate; \( \Delta \lambda := \bar{\lambda} - \Lambda \) is the difference in arrival rates; and \( j_+(\pi) := \frac{\pi}{\Lambda(\pi)} \) 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:

\[
\pi^3 := \frac{(c - \Delta) \mu}{(\bar{\lambda} - c)(\mu + 1) + (c - \Lambda) \mu},
\]

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

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

\[
(b) \quad \rho > \Delta \lambda \left( \frac{\bar{\lambda} - c}{2c - \bar{\lambda} - \Lambda} \right) + \Lambda \left( \frac{\Lambda}{\bar{\lambda}} \right) \left( \frac{\pi - c}{2c - \pi - \Delta} \right) - \Lambda.
\]

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

Conditions (a) and (b) reduce to the benchmark characterization of costly research provided in the article when \( \Lambda = 0 \).

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

**Proposition (OA2).** \( \lim_{\Lambda \to 0} \pi^3(\Lambda) \propto c - \rho \).

**Proof.** The Implicit Function Theorem implies that \( \mu(\Lambda) \) is differentiable, and \( \mu'(\Lambda) > 0 \). Thus, \( \pi^3(\Lambda) \) is also differentiable; we have:

\[
\pi'(\Lambda) = \frac{-(\bar{\lambda} - c) \mu(\Lambda) (1 + \mu(\Lambda)) + (\bar{\lambda} - c) (c - \Lambda) \mu'(\Lambda)}{[(\bar{\lambda} - c)(\mu(\Lambda) + 1) + (c - \Lambda) \mu(\Lambda)]^2}.
\]
By L'Hopital's rule, we have that $\lim_{\lambda \to 0} \mu'(\lambda) = \frac{1+\mu}{\lambda} = \frac{1+\mu(0)}{\lambda}$, and so:

$$
\lim_{\lambda \to 0} \pi^3(\lambda) = \frac{-\lambda(1-c)\mu(0)(1+\mu(0)) + (\lambda-c)c\lambda(1+\mu(0))}{[\lambda-c+\lambda\mu(0)]^2}
= \frac{(\lambda-c)(1+\mu(0))}{\lambda[\lambda-c+\lambda\mu(0)]^2}(c-\rho).
$$

The result follows.

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

$$(\rho + \lambda + \Delta)w(\pi) = \rho(\lambda + \Delta - 2c) + \Lambda(\pi)w(j+(\pi)) + \Lambda(1-\pi)w(j-(\pi)).$$

Let $\pi^4 > \frac{1}{2}$ 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 $1-\pi^4$, 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 $j_+(\pi^4)$ and $\pi^4$ will jump upwards to the region $(\pi^4, 1)$; similarly, posteriors in between $1-\pi^4$ and $j_-^1(1-\pi^4)$ will jump downward to the region $(0, 1-\pi^4)$. If the intervals $[1-\pi^4, j_-^1(1-\pi^4))$ and $(j_+^1(\pi^4), \pi^4]$ intersect, the value function on the intersection satisfies:

$$(\rho + \lambda + \Delta)w(\pi) = \rho(\lambda + \Delta - 2c) + \Lambda(\pi)[\Lambda(j_+(\pi)) - c + K(1-j_+(\pi))\Omega(j_+(\pi))^\mu] + \Lambda(1-\pi)[\Lambda(1-j_-(\pi)) - c + Kj_-(\pi)\Omega(1-j_-(\pi))^\mu].$$

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

$$(\rho + \lambda + \Delta)w(\pi) = \rho(\lambda + \Delta - 2c) + \Lambda(\pi)[\Lambda(j_+(\pi)) - c + K(1-j_+(\pi))\Omega(j_+(\pi))^\mu] + \Lambda(1-\pi)w(j_-(\pi)).$$

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 $1-\pi^4$. 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 $[1-\pi^4, \pi^4]$. The VM and SP conditions pin down the thresholds and constants of integration.
References
