

# 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:

$$w(\pi) = \max \left\{ 0, \bar{\lambda}\pi - c + \frac{\bar{\lambda}\pi (w(1) - w(\pi)) - \bar{\lambda}\pi(1 - \pi)w'(\pi)}{\rho}, \right. \\ \left. \bar{\lambda}(1 - \pi) - c + \frac{\bar{\lambda}(1 - \pi) (w(0) - w(\pi)) + \bar{\lambda}\pi(1 - \pi)w'(\pi)}{\rho}, \right. \\ \left. \bar{\lambda} - 2c + \frac{\bar{\lambda}(\bar{\lambda} - c - w(\pi))}{\rho} \right\}, \quad (1)$$

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

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

$$\text{(VM)} \quad w(\bar{\pi}) = w(1 - \bar{\pi}) =$$

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

$$\text{(SP)} \quad w'(\bar{\pi}) = w'(1 - \bar{\pi}) = 0.$$

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

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

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

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

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

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

the Bellman equation:

$$w(\pi) = \max \left\{ 0, \Lambda(\pi) - c + \frac{\Lambda(\pi)(w(j_+(\pi)) - w(\pi)) - \Delta\lambda\pi(1-\pi)w'(\pi)}{\rho}, \right. \\ \Lambda(1-\pi) - c + \frac{\Lambda(1-\pi)(w(j_-(\pi)) - w(\pi)) + \Delta\lambda\pi(1-\pi)w'(\pi)}{\rho}, \quad (2) \\ \left. \bar{\lambda} + \underline{\lambda} - 2c + \frac{\Lambda(\pi)w(j_+(\pi)) + \Lambda(1-\pi)w(j_-(\pi)) - (\bar{\lambda} + \underline{\lambda})w(\pi)}{\rho} \right\},$$

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

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

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

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

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

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

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  $\bar{\pi}^3 = \bar{\pi}^3(\underline{\lambda})$ ; notice that  $\bar{\pi}^3(0) = \bar{\pi}^1$ . All other parameters remain fixed throughout.

**Proposition (OA2).**  $\lim_{\underline{\lambda} \rightarrow 0} \bar{\pi}^{3'}(\underline{\lambda}) \propto c - \rho$ .

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

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

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

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

The result follows. □

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

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

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

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

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

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

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  $1 - \bar{\pi}^4$ . 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  $[1 - \bar{\pi}^4, \bar{\pi}^4]$ . 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.