Strategic Experimentation with asymmetric safe options*

Kaustav Das[†]   Nicolas Klein [‡]   Katharina Schmid[§]

June 29, 2023

Abstract

We study a two-player game of strategic experimentation with exponential bandits à la Keller, Rady and Cripps (2005) where the safe-arm payoff is different across players. We show that, as in Das, Klein and Schmid (2020), there exists an equilibrium in cutoff strategies if and only if the difference in safe-arm payoffs is large enough. In the equilibrium in cutoff strategies, the player with the higher safe-arm payoff conducts less experimentation. This feature of the equilibrium offers an explanation for the fact that oftentimes technological innovations are due to startups rather than established market leaders.

JEL Classification Numbers:C73, D83, O31 Keywords: Two-armed Bandit, Heterogeneous Agents, Free-Riding, Learning

1   Introduction

There are many situations when information generated through exploration by one agent is helpful to other agents as well. For example, when, in 2008, Tesla launched the Roadster, the first highway-compatible all-electric car, using lithium ion battery cells, other automobile companies became aware of the prospect of using lithium ion; hence gradually, electric cars were

---

launched by companies like Nissan, Citroën and BMW. In the economics literature, such games with informational externalities[1] have been analysed using models of strategic experimentation with bandits. Most of this literature, however, has been concerned with homogeneous players. However, car manufacturers clearly differ widely in terms of market shares. For example, when Tesla was launched in 2003, non-electric automobile markets were dominated by companies like General Motors and Toyota. In terms of a two-armed bandit model, this means the safe-arm payoffs were different across players. In this article, we aim to analyse the impact of asymmetric safe-arm options in a game of strategic experimentation with two-armed exponential bandits.

The strategic exponential-bandit model with homogeneous players has been introduced in a seminal paper by Keller, Rady and Cripps (2005). Das, Klein and Schmid (2020) has introduced asymmetry between the players in terms of different payoff arrival rates. In this article, we consider a variant of the two-armed bandit model where the asymmetry pertains to players' safe-arm payoffs. While we assume the payoff arrival rates on a good risky arm to be the same across the players, each player has a different flow payoff from their respective safe arm. In all other aspects, our model is identical to the canonical strategic exponential-bandit model of Keller, Rady and Cripps (2005). The latter show that with homogeneous players, there does not exist a Markov perfect equilibrium in which both players use a cut-off strategy. In any equilibrium, players swap the roles of pioneer and free-rider at least once. By contrast, Das, Klein and Schmid (2020) shows that, with sufficiently different payoff arrival rates from a good risky arm, there exists an equilibrium where both players use cut-off strategies. The equilibrium is unique in the class of equilibria in cutoff strategies; whenever only one player free-rides and the other experiments, it is always the player with the lower arrival rate who free-rides in the equilibrium in cutoff strategies.

In this article, we show that this feature is robust to the type of asymmetry across players. In particular, we show that, if the safe arm flow payoffs across players are sufficiently different, there exists an equilibrium where both players use a cut-off strategy. As in Das, Klein and Schmid (2020), this equilibrium is unique in the class of MPE in cut-off strategies. Whenever only one player experiments in equilibrium, it is the player with the lower safe-arm payoff.

In the context of our automobile example, this implies that a company that has a greater market share for non-electric vehicles will put relatively less effort into innovating by, e.g., offering an electric vehicle than a new entrant with a lower safe-arm payoff. It is indeed the case that the first highway-compatible electric vehicle was—arguably surprisingly—launched by a

---

[1]At the time of the introduction of a new product, a market will typically be very far from saturation, so that it seems reasonable to abstract from product-market competition between firms.

small startup company from California named Tesla, rather than by the established automobile giants like General Motors or Toyota. There are other instances as well when technological innovations have come from startups rather than established market leaders. For example, Christensen (2013) finds that over a period of two decades from 1976, each new generation of hard drives became smaller in size than the previous one, and a different set of companies dominated the market in each generation. In a recent paper, Awaya and Krishna (2021) have offered an explanation as to why, in many instances, technological innovations come from startups rather than firms that are already established in the respective industry. They consider a strategic bandit model where players can either innovate or not. The established incumbent is better informed about the viability of the innovation. Their main result shows that the incumbent tends to innovate less as better information dulls its incentive to learn from its rival. Our equilibrium in cutoff strategies offers a simple alternative explanation for the same phenomenon. Our model has the feature that if a company is working towards an innovation, it has to divert some resources from the market where it is already established. Hence, the opportunity cost of pursuing an innovation is higher for an already established market leader than for a startup; hence, the latter will be more inclined to persevere in its pursuit of innovation.

1.1 Related Literature

The paper contributes to the literature on strategic bandit models, a problem studied widely in economics. Some of the seminal papers in this area are by Bolton and Harris (1999), Keller, Rady and Cripps (2005), Keller and Rady (2010), Klein and Rady (2011), Klein (2013) and Thomas (2021). All these papers deal with symmetric players. The paper that is closest to the present article is Das, Klein and Schmid (2020), where we demonstrate how an equilibrium in cut-off strategies can exist if players have different learning abilities concerning the risky arm. In the current paper, we show that the same conclusion holds if the players have different safe-arm payoffs, while their innate exploration abilities are the same.

The rest of the paper is organised as follows. Section 2 describes the two-armed bandit model with different safe-arm payoffs. Sections 3 and 4 analyse the social planner's problem, and the non-cooperative game, respectively. Section 5 concludes. Our results are proved in the Appendix.

2 Two-armed bandit model with different safe-arm payoffs

There are two players (1 and 2), each of whom faces a two-armed bandit in continuous time. One of the arms is safe. If player $i \in \{1,2\}$ uses it, he gets a flow payoff of s_i, where $0 < s_1 < s_2$. The risky arm can be either good or bad. Both players' risky arms are of the same quality. If the risky arm is good, then a player using it receives a lump sum at the jumping times of a Poisson process with parameter $\lambda > 0$; the lump sum is drawn from a time-invariant distribution with mean $h > 0$. If the risky arm is bad, it never yields any payoff. Thus, a good risky arm gives both players an expected flow payoff of $g = \lambda h$, where we assume that $0 < s_1 < s_2 < g$. In all other aspects, the model is similar to the one in Das, Klein and Schmid (2020). The parameter values and the game are common knowledge.

Players do not initially know whether their risky arms are good or bad. They start with a common prior belief $p_0 \in (0,1)$ that their risky arms are good. Players have to decide in continuous time whether to choose the risky arm or the safe arm, and at each instant a player can choose only one arm. We write $k_{i,t} = 1(0)$ if player $i \in \{1,2\}$ uses the risky (safe) arm at instant $t \in \mathbb{R}_+$. Players' actions and outcomes are perfectly publicly observable. This implies that players have a common posterior belief p_t at all times $t \geq 0$. Players discount the future according to the common discount rate $r > 0$.

Let p_t be the players' common belief that their risky arm is good at time $t \geq 0$. Since only a good risky arm can ever yield positive payoffs, the arrival of a lump sum fully reveals the risky arm to be good. In the absence of a lump-sum arrival, by Bayes' rule, the belief follows the following law of motion:

$$dp_t = -\lambda(k_{1,t} + k_{2,t})p_t(1 - p_t)\,dt.$$

In the next section, we characterise a utilitarian planner's solution.

3 The Planner's Solution

The planner's objective is to maximise the sum of the players' expected discounted payoffs. The planner's action is denoted by the pair (k_1, k_2) $(k_i \in \{1,2\})$, where $k_i = 0(1)$ denotes that the planner has allocated player i to the safe (risky) arm. The planner's value function v satisfies the Bellman equation

$$v(p) = s_1 + s_2 + \max_{k_1 \in \{0,1\}} k_1[B(p,v) - c_1(p)] + \max_{k_2 \in \{0,1\}} k_2[B(p,v) - c_2(p)]$$

where $B(p,v) = \frac{\lambda p\{2g - v - v'(1-p)\}}{r}$ and $c_i(p) = s_i - gp$ $(i = 1,2)$. We characterise the planner's solution in the following proposition.

Proposition 1 Let $p_1^* := \frac{rs_1}{(2\lambda + r)g - \lambda(s_1 + s_2)}$. There exists a threshold $p_2^* \in \left(p_1^*, \frac{s_2}{g}\right)$ such that the planner's optimal policy is given by

$$(k_1^*(p), k_2^*(p)) = \begin{cases} (1,1) & \text{if } p > p_2^* \\ (1,0) & \text{if } p_1^* < p \le p_2^* \\ (0,0) & \text{if } p \le p_1^* \end{cases}.$$

The planner's value is C^1, (strictly) increasing and convex (on $(p_1^*, 1)$), and is given by

$$v(p) = \begin{cases} 2pg + C_{11}(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{2\lambda}} & =: v^{(1,1)}(p) & \text{if } p > p_2^* \\ s_2 + \frac{(2\lambda+r)g - \lambda s_2}{\lambda+r}p + \frac{1-p}{1-p_1^*}\left(\frac{(1-p)p_1^*}{(1-p_1^*)p}\right)^{\frac{r}{\lambda}}\left[s_1 - \frac{(2\lambda+r)g - \lambda s_2}{\lambda+r}p_1^*\right] & =: v^{(1,0)}(p) & \text{if } p_1^* < p \le p_2^* \\ s_1 + s_2 & =: v^{(0,0)}(p) & \text{if } p \le p_1^* \end{cases},$$

where the threshold p_2^* is implicitly defined by $v^{(1,0)}(p_2^*) = 2s_2$, and the constant of integration $C_{11} > 0$ is such that $v^{(1,1)}(p_2^*) = 2s_2$.

The proof can be found in Appendix (A). Next, we turn our attention to the analysis of the non-cooperative game.

4 The Non-Cooperative Game

In the non-cooperative game, we restrict ourselves to Markovian strategies with the common posterior belief as the state variable, defined as mappings $k_i : [0,1] \to \{0,1\}$, where $k_i(p) = 0(1)$ indicates that player $i \in \{1,2\}$ is choosing the safe (risky) arm, if the common posterior belief is $p \in [0,1]$. The Bellman equations for players 1 and 2's value functions can respectively be written as

$$v_1(p) = s_1 + k_2 b(p,v_1) + \max_{k_1 \in \{0,1\}} k_1[b(p,v_1) - c_1(p)],$$

and

$$v_2(p) = s_2 + k_1 b(p, v_2) + \max_{k_2 \in \{0,1\}} k_2 [b(p, v_2) - c_2(p)],$$

where $b(p, v) = \frac{\lambda p \{g - v - v'(1-p)\}}{r}$.

We shall now determine the players' best responses, using the same method as Keller, Rady and Cripps (2005) or Das, Klein and Schmid (2020). Suppose player 2 is choosing the safe arm in some neighbourhood of $p \in (0,1)$. Then, player 1 chooses the risky arm at p if and only if the belief p exceeds his single-agent threshold $\bar{p}_1 = \frac{rs_1}{(r+\lambda)g - \lambda s_1}$. By the same token, if player 1 is choosing the safe arm in a neighbourhood of p, player 2 will choose the risky arm as long as the belief is greater than $\bar{p}_2 := \frac{rs_2}{(r+\lambda)g - \lambda s_2}$. Since, $s_1 < s_2$, we have $\bar{p}_1 < \bar{p}_2$.

Now, suppose that player 2 is playing risky in some neighbourhood of $p \in (0,1)$. Player 1's best response is to choose the risky arm if and only if $b(p, v_1) \geq c_1(p)$, and hence $v_1 \geq s_1 + c_1(p)$, in this neighbourhood. This means that given the other player is choosing the risky arm, choosing the risky arm constitutes a best response for player 1 if and only if v_1 lies above D_1 in the (p, v)-plane, where the diagonal D_1 is defined as

$$D_1 := \{(p, v) \in [0, 1] \times \mathbb{R}_+ : v = s_1 + c_1(p)\}.$$

Similarly, for player 2, given that player 1 is choosing the risky arm, choosing the risky arm constitutes a best response for player 2 if and only if v_2 is above D_2 in the (p, v)-plane,[2] where

$$D_2 := \{(p, v) \in [0, 1] \times \mathbb{R}_+ : v = s_2 + c_2(p)\}.$$

The players' best-response diagonals are depicted in Figure 1.

From Figure 1, it can be observed that heterogeneity in safe-arm payoffs drives a wedge between players' best response lines. When $s_1 = s_2$, D_1 and D_2 coincide, while, when $s_2 = g$, the lines are the farthest apart from each other. The next proposition shows that if and only if players are sufficiently heterogeneous, there exists an equilibrium where both players use cutoff strategies; moreover, this equilibrium is unique in the class of equilibria where both players use cutoff strategies.

Proposition 2 There exists a $s_2^* \in (s_1, g)$ such that there exists an equilibrium where both players

---

[2]In a slight abuse of notation, we shall in the following also write D_i ($i \in \{1, 2\}$) for the function $\tilde{D}_i : [0, 1] \to \mathbb{R}$ such that D_i be the graph of $\tilde{D}_i$, i.e., $(p, \tilde{D}_i(p)) \in D_i$ for all $p \in [0, 1]$.
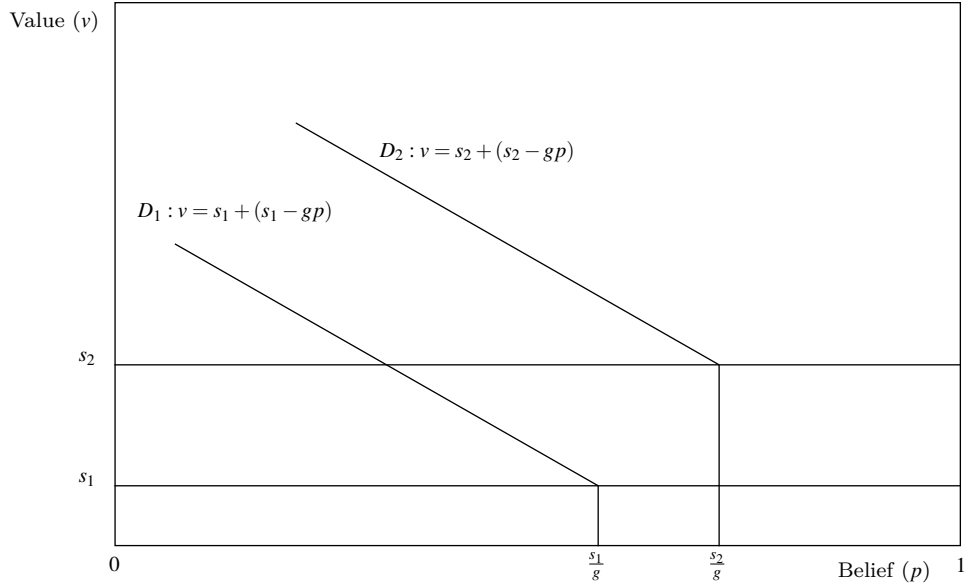
Figure 1.

use cutoff strategies if and only if $s_2 \in [s_2^*, g)$. In this equilibrium, player 1 chooses the risky arm for all beliefs greater than $\bar{p}_1$ and chooses the safe arm otherwise. There exists a unique $p_2' \in (\bar{p}_1, 1)$ such that player 2 chooses the risky arm for beliefs greater than p_2' and the safe arm otherwise. This equilibrium is unique in the class of equilibria where both players use cutoff strategies.

The formal proof of this proposition is given in Appendix (B). It can be seen from Figure 2 that, as players' safe-arm payoffs diverge from each other, the range of beliefs over which only player 2 can free-ride expands. Just as in Das, Klein and Schmid (2020), this difference in free-riding opportunities allows for the existence of an equilibrium in cutoff strategies. The equilibrium is depicted in Figure 2. The equilibrium in cutoff strategies exists only if the curve v_1 intersects D_1 at a belief that is lower than the one at which the curve v_2 intersects D_2. This is possible only if the players are sufficiently heterogeneous.

Thus, as in Das, Klein and Schmid (2020), sufficient heterogeneity, and hence sufficiently different free-riding opportunities, allow for the existence of a cutoff equilibrium. A larger difference in players' safe-arm payoffs reduces the free-riding opportunities of player 1, while expanding those of player 2.
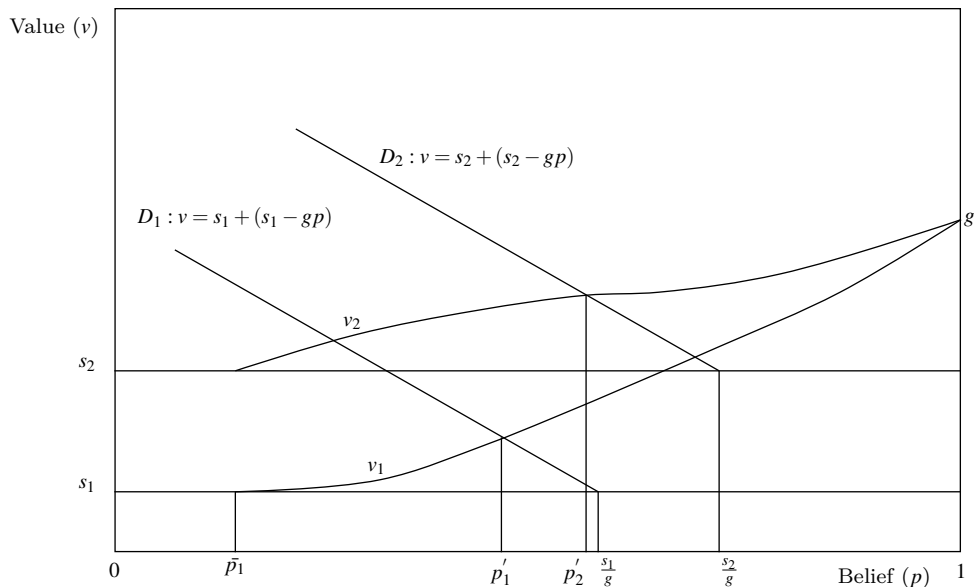
7

Figure 2.

5   Conclusion

In this article, we have considered a two-armed strategic bandit model where players face different opportunity costs of experimentation in terms of safe-arm payoffs but an identical rate of learning on the risky arm. We show that, as in Das, Klein and Schmid (2020), there exists a Markov Perfect equilibrium in cut-off strategies if and only if players are sufficiently asymmetric, i.e., the difference in safe-arm payoffs is high enough. Hence, we show that the conclusion that a sufficient degree of asymmetry across players is necessary and sufficient for the existence of an equilibrium in cutoff strategies continues to apply also when the asymmetry pertains to the opportunity costs, rather than the yield, of information acquisition via the risky arm. Furthermore, our equilibrium in cut-off strategies offers a simple explanation for the observation that, in many instances, innovation will come from start-ups rather than established companies, the idea being that start-ups have less to lose by focussing their resources and energy on the pursuit of an innovation.

References

Awaya, Y. and Krishna, V., 2021. Startups and upstarts: disadvantageous information in r&d. Journal of Political Economy, 129(2), pp.534-569.

Bolton, P. and Harris, C., 1999. Strategic experimentation. Econometrica, 67(2), pp.349-374.

Das, K., Klein, N. and Schmid, K., 2020. Strategic experimentation with asymmetric players. Economic Theory, 69, pp.1147-1175.

Christensen, C.M., 2013. The innovator's dilemma: when new technologies cause great firms to fail. Harvard Business Review Press.

Keller, G., Rady, S. and Cripps, M., 2005. Strategic experimentation with exponential bandits. Econometrica, 73(1), pp.39-68.

Keller, G. and Rady, S., 2010. Strategic experimentation with Poisson bandits. Theoretical Economics, 5(2), pp.275-311.

Klein, N., 2013. Strategic learning in teams. Games and Economic Behavior, 82, pp.636-657.

Klein, N. and Rady, S., 2011. Negatively correlated bandits. The Review of Economic Studies, 78(2), pp.693-732.

D. Thomas, C., 2021. Strategic experimentation with congestion. American Economic Journal: Microeconomics, 13(1), pp.1-82.

APPENDIX

A   Proof of Proposition 1

We prove the propoasition in two steps. First, we compute the planner's payoff from the proposed policy, and show that it is C^1. Then, we show that this payoff and the associated policy solve the Bellman equation.

First, consider the range $p \in [0, p_1^*]$. According to the conjectured solution, $k_1^*(p) = k_2^*(p) = 0$, and $v(p) = s_1 + s_2$.

For $p \in (p_1^*, p_2^*]$, according to the conjectured solution, $k_1^*(p) = 1$ and $k_2^*(p) = 0$. This implies that $v(p)$ satisfies the ODE

$$\lambda p(1-p)v'(p) + (r + \lambda p)v(p) = rs_2 + (r + 2\lambda)pg,$$

which is solved by

$$v^{(1,0)}(p) = s_2 + \frac{(2\lambda + r)g - \lambda s_2}{\lambda + r}p + C_{10}(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda}},$$

9

where $C_{10} = \frac{1}{1-p_1^*}\left(\frac{1-p_1^*}{p_1^*}\right)^{-\frac{r}{\lambda}}\left[s_1 - \frac{(2\lambda+r)g-\lambda s_2}{\lambda+r}p_1^*\right] > 0$ is a constant of integration, which is determined by value matching at $p = p_1^*$; direct computation verifies that smooth pasting obtains as well, i.e., that $v^{(1,0)'}(p_1^*) = 0$. As $C_{10} > 0$, $v^{(1,0)}$ is strictly convex and hence strictly increasing. Moreover, $v^{(1,0)}(1) > 2s_2 > s_1 + s_2 = v^{(1,0)}(p_1^*)$. Thus, $p_2^* \in (p_1^*, 1)$ is well-defined.

For $p > p_2^*$, both players use the risky arm according to the conjectured solution. Hence, the planner's value $v(p)$ satisfies the ODE

$$2\lambda p(1-p)v'(p) + (r+2\lambda p)v(p) = 2(r+2\lambda)pg,$$

which is solved by

$$v^{(1,1)}(p) = 2pg + C_{11}(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{2\lambda}},$$

where C_{11} is a constant of integration determined by $v^{(1,1)}(p_2^*) = 2s_2 = v^{(1,0)}(p_2^*)$. It is immediate to verify from the appertaining ODEs that this implies that $v^{(1,1)'}(p_2^*) = v^{(1,0)'}(p_2^*)$ (smooth pasting), and that therefore v is strictly increasing and convex on $(p_2^*, 1)$ as well. As is immediate from the expression for $v^{(1,1)}$, $C_{11} > 0$ if and only if $p_2^* < \frac{s_2}{g}$. To show that indeed $p_2^* < \frac{s_2}{g}$, it is sufficient to note that $v^{(1,0)}\left(\frac{s_2}{g}\right) > 2s_2$, which is easily verified as $v^{(1,0)}\left(\frac{s_2}{g}\right) - C_{10}\left(1 - \frac{s_2}{g}\right)\left(\frac{g-s_2}{s_2}\right)^{\frac{r}{\lambda}} > 2s_2$.

It remains to show that the proposed policy satisfies the planner's Bellman equation. First, on $[0, p_1^*]$, $v(p) = s_1 + s_2$, and therefore $B(p,v) = \frac{\lambda}{r}p(2g - s_1 - s_2) \le s_i - pg = c_i(p)$, as $p \le p_1^*$.

On (p_1^*, p_2^*), $B(p,v) = v(p) - pg - s_2$, so that $B(p,v) \ge c_1(p)$ if and only if $v(p) \ge s_1 + s_2$, and $B(p,v) \le c_2(p)$ if and only if $v(p) \le 2s_2$; both inequalities hold by the monotonicity of v.

On $(p_2^*, 1)$, $B(p,v) = \frac{v(p)}{2} - pg$, and thus $B(p,v) \ge c_2(p) \ge c_1(p)$ if and only if $v(p) \ge 2s_2$, which holds by the monotonicity of v. ∎

B   Proof of Proposition 2

The same argument as in Das, Klein and Schmid (2020), Proposition 2, establishes that neither player will play risky on $[0, \bar{p}_1]$ in any equilibrium.

In any equilibrium, for beliefs just above $\bar{p}_1$, only player 1 can experiment with the risky arm, as $\bar{p}_1 < \bar{p}_2$ and the point $(\bar{p}_1, s_i)$ lies below both diagonals D_1 and D_2 in (p,v)-space. This implies that, for beliefs just above $\bar{p}_1$, the payoff of player 1 follows the ODE

$$\lambda p(1-p)v'(p) + (r+\lambda p)v(p) = (r+\lambda)pg,$$

which is solved by

$$v_1^{(1,0)}(p) = gp + C_1^{(1,0)}(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda}}.$$

The integration constant $C_1^{(1,0)}$ is determined by the value-matching condition $v_1^{(1,0)}(\bar{p}_1) = s_1$.

On an open interval of beliefs in which only player 1 is experimenting, player 2's payoff satisfies the ODE

$$\lambda p(1-p)v'(p) + (r+\lambda p)v(p) = rs_2 + \lambda pg,$$

which is solved by

$$v_2^{(1,0)}(p) = s_2 + \frac{\lambda}{\lambda+r}(g-s_2)p + C_2^{(1,0)}(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda}}.$$

The integration constant $C_2^{(1,0)}$ is determined by the value-matching condition $v_2^{(1,0)}(\bar{p}_1) = s_2$.

First, we show that there exists a unique $p_1' \in (\bar{p}_1, 1)$ such that $v_1^{(1,0)}(p_1') = D_1(p_1')$. Indeed, as $\bar{p}_1 g < s_1$, $C_1^{(1,0)} > 0$, and $v_1^{(1,0)}$ is strictly convex. It is easy to check that $v_1^{(1,0)'}(\bar{p}_1) = 0$, i.e., smooth pasting holds at $\bar{p}_1$. Therefore, $v_1^{(1,0)}$ is strictly increasing on $(\bar{p}_1, 1)$. The diagonal D_1, on the other hand, is monotonically decreasing in p. Moreover, since $\bar{p}_1 < \frac{s_1}{g}$, we have $s_1 = v_1^{(1,0)}(\bar{p}_1) < D_1(\bar{p}_1)$, and $g = v_1^{(1,0)}(1) > D_1(1) = 2s_1 - g$. Hence, there exists a unique $p_1' \in (\bar{p}_1, 1)$ such that $v_1^{(1,0)}(p_1') = D_1(p_1')$. Note that since both $v_1^{(1,0)}$ and D_1 are independent of s_2, the threshold p_1' is independent of s_2.

Next, we will argue that there exists a unique $p_2' \in (\bar{p}_1, 1)$ such that $v_2^{(1,0)}(p_2') = D_2(p_2')$. As $s_2 + \frac{\lambda}{\lambda+r}(g-s_2)\bar{p}_1 > s_2$, $C_2^{(1,0)} < 0$, and $v_2^{(1,0)}$ is strictly concave; it is moreover strictly increasing. The diagonal D_2 is monotonically decreasing in p. Moreover, since $\bar{p}_1 < \frac{s_1}{g} < \frac{s_2}{g}$, we have $s_2 = v_2^{(1,0)}(\bar{p}_1) < D_2(\bar{p}_1)$, and $\frac{\lambda g + rs_2}{\lambda+r} = v_2^{(1,0)}(1) > D_2(1) = 2s_2 - g$. Thus, there exists a unique $p_2' \in (\bar{p}_1, 1)$ such that $v_2^{(1,0)}(p_2') = D_2(p_2')$.

By the properties of the best-response correspondences, an equilibrium in cutoff strategies can exist only if $p_2' \geq p_1'$. When $s_2 \uparrow g$, $p_2' \to 1$. Moreover, $v_2^{(1,0)}(p) \approx s_2$ for s_2 close to g. Since $p_1' \in (\bar{p}_1, 1)$ is independent of s_2, $p_2' \approx 1 > p_1'$ for s_2 close enough to g. For the case $s_2 = s_1$, we know by Keller, Rady and Cripps (2005) that $p_2' < p_1'$. As s_2 increases, the line D_2 shifts upward and to the right, while $v_2^{(1,0)}$, a strictly increasing function satisfying $v_2^{(1,0)}(\bar{p}_1) = s_2 < \frac{\lambda g + rs_2}{\lambda+r} = v_2^{(1,0)}(1)$ becomes flatter. (Recall that $\bar{p}_1$ is independent of s_2.) Hence, p_2' increases as s_2 increases. Thus, there exists a unique $s_2^* \in (s_1, g)$ such that, for all $s_2 \in [s_2^*, g)$, $p_2' \geq p_1'$.

We have shown so far that $s_2 \geq s_2^*$ is necessary for the existence of an equilibrium in cutoff

strategies, and that the only candidate for an equilibrium in cutoff strategies is that in which player i applies the cutoff p_i'. It remains to show that these cutoffs indeed define an equilibrium whenever $s_2 \geq s_2^*$, i.e., that $k_1 = \mathbb{1}_{(p_1',1]}$ and $k_2 = \mathbb{1}_{(p_2',1]}$ are mutually best responses. For beliefs $p \in [0, p_2']$, this is so by construction. Let us therefore turn to the belief range $(p_2', 1]$.

In this range, the proposed strategies imply that player i's value function follows the ODE

$$2\lambda p(1-p)v_i'(p) + (r + 2\lambda p)v_i(p) = (r + 2\lambda)pg,$$

which is solved by

$$v_i^{(1,1)}(p) = pg + C_i^{(1,1)}(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{2\lambda}},$$

where $C_i^{(1,1)}$ is a constant of integration determined by $v_i^{(1,1)}(p_2') = v_i^{(1,0)}(p_2')$.

To show that player 1 plays a best response in this range, it suffices to show that $v_1^{(1,1)} \geq D_1$ throughout this range. As $p_2' \geq p_1'$, the monotonicity of $v_1^{(1,0)}$ and D_1 imply that $v_1^{(1,0)}(p_2') = v_1^{(1,1)}(p_2') \geq D_1(p_2')$. Now suppose to the contrary that there exists some $\hat{p} \in [p_2', 1)$ such that $v_1^{(1,1)}(\hat{p}) = D_1(\hat{p})$ and $v_1^{(1,1)'}(\hat{p}) \leq -g = D_1'(\hat{p})$. By the ODE for $v_1^{(1,1)}$, this implies $\hat{p} \leq \frac{rs_1}{(2\lambda + r)(g - s_1) + rs_1} < \bar{p}_1 < p_1' \leq p_2' \leq \hat{p}$, a contradiction. Thus, $v_1^{(1,1)}(p) \geq D_1(p)$ for all $p \in [p_2', 1]$, and player 1 is playing a best response in this range as well.

Now let us turn to player 2. The payoff function $v_2^{(1,1)}$ is convex if and only if $C_2^{(1,1)} \geq 0$, i.e., if and only if $v_2^{(1,1)}(p_2') = v_2^{(1,0)}(p_2') = D_2(p_2') = 2s_2 - p_2'g \geq p_2'g$, which is equivalent to $p_2' \leq \frac{s_2}{g}$; this holds because $v_2^{(1,0)}$ is strictly increasing, D_2 is strictly decreasing, $s_2 = v_2^{(1,0)}(\bar{p}_1) = D_2(\frac{s_2}{g})$, and $\bar{p}_1 < \frac{s_2}{g}$. Thus, $v_2^{(1,1)}$ is convex. Direct computation from the appertaining ODEs furthermore shows smooth pasting at p_2', i.e., $v_2^{(1,1)'}(p_2') = v_2^{(1,0)'}(p_2') > 0$. We can thus conclude that $v_2^{(1,1)}$ is monotonically increasing on $(p_2', 1)$, and therefore $v_2^{(1,1)}(p) > D_2(p)$ for all $p \in (p_2', 1]$, as D_2 is monotonically decreasing. This shows that player 2 is playing a best response on $[p_2', 1]$ as well, thus concluding the proof. ∎