# Inefficient Duplication of Efforts in Patent Races[1]

Kaustav Das[2] Nicolas Klein[3]

July 4, 2019

[1]This paper supersedes an earlier version entitled "Competition, Learning and Duplicative Search in a Patent Race." We thank Siddhartha Bandyopadhyay, Dirk Bergemann, Kalyan Chatterjee, Martin Cripps, Sanjeev Goyal, Johannes Hörner, Vijay Krishna, Sven Rady, Larry Samuleson and Caroline Thomas for their helpful comments. Thanks are also due to the participants of the Warwick Lunchtime Theory Workshop 2019, the Game Theory Conference at Edinburgh 2019, the Game Theory Symposium at the University of York 2018, the Annual Conference at the Indian Statistical Institute New Delhi 2018, the North American Summer Meeting of the Econometric Society 2014 in Minneapolis, the European Meeting of the Econometric Society 2014 in Toulouse.

[2]Department of Economics, University of Exeter Business School; email: daskaustav84@gmail.com

[3]Université de Montréal and CIREQ, Email:kleinnic@yahoo.com

## Abstract

We analyse a model of two firms locked in a winner-takes-all competition. Firms have to choose in continuous time between a *traditional* and an *innovative* method of pursuing the decisive breakthrough. They share a common belief about the likelihood of the innovative method being good. The unique Markov perfect equilibrium is efficient if and only if firms are symmetric in their ability of leveraging a good innovative method. Otherwise, equilibrium will entail inefficient duplication of efforts in the innovative method. Inefficiency is worst for intermediate degrees of heterogeneity, and is mitigated if early completion of the project is promoted.

**JEL Classification Numbers:** C73, D83, O31.

**Keywords:** R&D competition, Duplication, Two-armed Bandit, Learning

# 1 Introduction

Process innovation is an important driver of success in many industries. Consider the pharmaceutical industry and its quest for a better way of treating Alzheimer's disease, for instance. Alzheimer's is characterised by both a decrease in acetylcholine (neurotransmitter) levels in the brain and the accumulation of $\beta$-amyloid plaques. The current method of treatment is based on the widely marketed drug Donepezil, which increases acetylcholine levels but which can only slow down the progression of the disease without curing it. Research efforts over the past decade, by contrast, have been focussed on finding a drug counter-acting the accumulation of $\beta$-amyloid plaques. As innovative approaches toward this goal have failed to lead to success, researchers are currently exploring the possibility of designing a drug that would combat the accumulation of $\beta$-amyloid plaques via an increase in neurotransmitter levels.[1] Indeed, there is some evidence that Donepezil has a beneficial effect on the level of $\beta$-amyloid plaques.[2]

When firms search for success using an innovative method, their competition entails a positive informational externality, besides the payoff externality that is typical for patent races. Indeed, the fact that a competitor has been unsuccessfully looking for a breakthrough using a particular method is useful information to the firm, as it will inform its future optimal R & D choices. In our Alzheimer's example, failed clinical trials by a pharmaceutical company indeed provide crucial insights that also help shape competitors' future research efforts.

In this paper, we study process innovation in a setting in which two firms are engaged in a patent race and their research choices are observable. There is an established work method either firm can use, which leads to a success at the first jumping time of a Poisson process with a known rate. As Donepezil is already known to have an effect on $\beta$-amyloid plaques, this would correspond to the search for a drug that seeks to fight the concentration of $\beta$-amyloid plaques by increasing

[1] See Moss (2018).

[2] See Dong et.al (2009).

neurotransmitter levels. Both firms also have access to an innovative work method that is either good or bad. Whether it is good or bad is initially unknown to the firms, who share a common initial belief about it. If the innovative method is good, it leads to a success at a faster rate than the established method. We allow for one firm to be more efficient than the other in its exploration of the innovative method, achieving a success more quickly conditionally on the method being good. If the innovative method is bad, it never yields a success. The first success yields a payoff only to the firm that produced it and ends the game. Both firms discount future payoffs at a common rate.

The innovative method is good for one firm if and only if it is good for the other firm as well. As either firm's actions are perfectly publicly observable, the R & D race between the two firms involves a positive informational externality. Indeed, the longer the innovative method is unsuccessfully tried by *either* firm, the more pessimistic *both* firms become about its quality.

We show that our game admits a unique Markov perfect equilibrium, with the firms' common belief that the innovative method is good as the state variable. In contrast to the case of pure informational externalities (see Keller, Rady and Cripps (2005)), our unique equilibrium is always in cutoff strategies. This is because unlike in Keller, Rady and Cripps (2005), free-riding is not an issue in our setting due to the winner-takes-all feature of the R & D race. If and only if both firms are equally productive with a good innovative work method the unique Markov perfect equilibrium is efficient.

By contrast, if one of the firms happens to be more productive with a good innovative method, e.g., because it has a bigger or better research or production department, the unique Markov perfect equilibrium leads to a welfare loss on account of inefficient duplication of innovative efforts. The stronger firm, however, always acts efficiently, which, in the case of our model, coincides with myopic behavior. Indeed, the stronger firm always gives up last in equilibrium. At beliefs just above its cutoff, therefore, its situation is that of a single agent, implying that it will act efficiently. As the game ends at the time of the first breakthrough, information is of no use after a success; a single agent will thus, at each point in time, optimally maximize the expected breakthrough rate, i.e., behave myopically.

The situation is more complicated for the less productive firm, which anticipates that the more productive firm will continue exploring the innovative method until its lower myopic cutoff is reached. At its own myopic cutoff, the less productive firm thus reasons that, if it goes on experimenting a bit longer, the more productive firm's myopic cutoff is reached sooner (conditionally on no success); put differently, the amount of time the productive firm will henceforth spend on exploring the innovative method is reduced. Based on the current belief, this means that the overall likelihood of a success by the more productive firm decreases. Since, at the less productive firm's myopic cutoff, its own expected breakthrough rate is exactly equalized between the two methods, this explains why a utilitarian planner will apply a cutoff more optimistic than its myopic cutoff to the less productive firm. Because of the payoff externality, the same reasoning explains why the less productive firm will extend experimentation below its myopic cutoff in equilibrium, leading to inefficient duplication of innovative efforts. In summary, the weaker firm is asked to step aside somewhat for the stronger firm in the planner's solution; in equilibrium, by contrast, the weaker firm explicitly endeavours to "eat up" some of the stronger firm's comparative advantage.

Pfizer has pulled out from Alzheimer's drug research in January 2018 while its competitors keep pursuing it, which suggests that heterogeneity among firms is indeed a feature of real-world R & D races. Analysing a large database that contains information on R & D projects for more than $28,000$ cases, Pammoli, Magazzini and Riccaboni (2011) conclude that, in the period they study, there has been a decline in R & D productivity in pharmaceuticals, which cannot be fully explained by the market forces of demand and competition. Simultaneously, they observe an increasing concentration of R & D investments in relatively more risky areas. Our model identifies a novel effect generating inefficient duplication of efforts in an R & D race with observable actions as a result of firm heterogeneity.

**Related Literature:** The problem of incentivizing a single agent to engage in innovation has been analysed by Klein (2016) in continuous time and by Manso (2011) in a two-period model. Of this latter setting, Ederer (2013) studies an extension to two agents.

Our model builds on the literature on strategic experimentation with bandits, started by Bolton and Harris (1999). In particular, we use a variant of the exponential model of Keller, Rady and Cripps (2005). Das, Klein and Schmid (2019) have introduced heterogeneity in Poisson arrival rates into this model. The negative payoff externality here is as in the treasure-hunt game of Chatterjee and Evans (2004), which is the first paper to analyse project choice in a dynamic winner-takes-all competition. In contrast to our setting, Chatterjee and Evans (2004)' players choose between two different work methods, exactly one of which is known to work, while players are initially uncertain which one it is. Whereas they do not analyze the impact of heterogeneity in players' abilities, they also find an efficient equilibrium when both research avenues imply the same cost. For the case of costs that are asymmetric across research avenues they show that there is either *too much* or *too little* exploration of a research avenue, depending on the players' prior belief.

Another closely related paper is Besanko and Wu (2013), who also introduce payoff externalities into a Keller, Rady and Cripps (2005) setting. Their paper differs from ours in that their safe option consists of an alternative *project* rather than of an alternative *process* for a given project; i.e., as in Keller, Rady and Cripps (2005), their safe project gives players an immediate known payoff. This difference in the nature of the safe option leads to a sharp difference in results: While our unique Markov perfect equilibrium, which is symmetric, is efficient for homogeneous players, Besanko and Wu (2013)'s unique symmetric Markov perfect equilibrium features over-experimentation in the case of a negative payoff externality.

Our paper also contributes to the relatively less explored area of choice of methodological approach in R & D races.[3] Indeed, we show that, when firms are heterogeneous, there is always some inefficient duplication of research effort. In a static model with winner-takes-all competition, Bhattacharya and Mookerjee (1986) show that, when firms are symmetric and not excessively risk-averse, market allocations and socially optimal allocations coincide, both requiring extreme specialisation. However, with sufficient risk aversion, there is a tendency towards

[3]The current paper is not concerned about the *scale* of R & D. We analyse the problem of allocation of a *given* resource among the various methods of R & D. The issue of choosing the scale of R & D is well documented in the literature (see Lee and Wilde (1980); Reinganum (1982)).

under duplication. Dasgupta and Maskin (1987), by contrast, assume that a project is the costlier the more unusual it is, and find that market research portfolios consist of projects that are too highly correlated. Letina (2016) analyses a static model where $N$ symmetric firms compete in the pre-innovation market by choosing a subset from a continuum of heterogeneous research projects. All approaches are initially equally likely to succeed and it is known that exactly one of them will. There is duplicative equilibrium effort in projects with lower costs, with fewer firms developing the more expensive approaches.

The rest of the paper is organised as follows. Section 2 describes the environment, while Section 3 and 4 describe the analysis with symmetric and asymmetric players respectively, while Section 5 discusses comparative statics. Section 6 concludes. Most proofs are in the Appendix.

# 2 The Environment

Two firms are simultaneously trying to be the first to achieve a breakthrough in continuous time. The first breakthrough yields a payoff of 1 to the firm accomplishing it; i.e., the first firm to innovate appropriates all the rent. There are two work methods the firms can adopt to achieve a breakthrough. One method is *established* (*non-risky*) in that it yields a breakthrough at the first jumping time of a Poisson process with known intensity $\lambda_0 > 0$. The other method is *innovative* (*risky*), in that it is not initially known if it is good or bad, its quality being the same for both firms. If it is good, it produces a breakthrough for firm $i \in \{1,2\}$ at the first jumping time of a Poisson process with intensity $\lambda_i > \lambda_0$. If it is bad, it never yields a breakthrough for either firm. We assume $\lambda_1 \geq \lambda_2$; i.e., conditionally on the innovative method being good, firm 1 will achieve the breakthrough weakly faster in expectation. For the rest of the paper, the *established* method will be denoted $S$ and the *innovative* method will be denoted $R$. Both firms discount the future using the common discount rate $r > 0$. Firms do not incur any direct costs for adopting either method. They share a common prior $p \in (0,1)$ that method $R$ is good. Firms' choices of methods are perfectly publicly observable. This implies that, at any time point in time, firms will also share a common posterior belief.

**Evolution of beliefs:** If $k_{i,t}$ is an indicator variable for firm $i$ adopting the innovative method, then conditionally on no success arriving via the innovative method, the common posterior $p_t$ evolves a.s. according to

$$dp_t = -(k_{1,t}\lambda_1 + k_{2,t}\lambda_2)p_t(1-p_t)\,dt.$$

# 3 Symmetric Firms

In this section, we analyse the case of firms that are symmetric in their ability to achieve a success by a good innovative method. This means we have $\lambda_1 = \lambda_2 > \lambda_0$. We first analyse the social planner's problem, who seeks to maximise the firms' average discounted payoffs.

## 3.1 Social Planner's problem

Without loss of generality we can restrict the planner to Markov strategies $k(p_t)$ with the posterior belief $p_t$ as the state variable, where $k$ denotes the number of firms the planner assigns to method $R$.[4] This implies $k(p_t) \in \{0,1,2\}$. Let $v(p)$ be the value function of the planner. Then we have

$$rv = \max_{k\in\{0,1,2\}}\{\lambda_0(1-2v) + k[\lambda_1 p\left(\frac{1}{2} - v - v^{'}\cdot(1-p)\right) - \lambda_0\left(\frac{1}{2} - v\right)]\}$$

The expression $\lambda_0(1-2v)$ denotes the expected flow payoff the planner can guarantee himself by using the method S. The expression $\lambda_1 p\left(\frac{1}{2} - v - v^{'}(1-p)\right) - \lambda_0\left(\frac{1}{2} - v\right)$ reflects the premium the planner gets by assigning an additional firm to method $R$. Note that, by linearity, even if firms' efforts were divisible, it would be without loss for the planner to choose $\{k(p_t)\}_{t\geq 0}$ with $k(p_t) \in \{0,2\}$ for all $t \geq 0$. The planner's solution is described in the following proposition. It shows that efficiency requires players to choose the myopically optimal method. To state the proposition, we use the function $\mu(p)$ defined in Appendix A. In the homogeneous case, $\mu(p) = (1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{2\lambda_1}}$.

[4] We suppress the arguments, whenever this is convenient.

**Proposition 1** *The planner's optimal policy $k^*(p)$ is given by*

$$k^*(p) = \begin{cases} 2 & \text{if } p \in (p_1^*, 1] \\ 0 & \text{if } p \in [0, p_1^*] \end{cases},$$

*where $p_1^* = \frac{\lambda_0}{\lambda_1}$. The planner's value function is given by*

$$v(p) = \begin{cases} \frac{\lambda_1}{r+2\lambda_1} p + [\frac{\lambda_0}{r+2\lambda_0} - \frac{\lambda_1}{r+2\lambda_1} p_1^*] \frac{\mu(p)}{\mu(p_1^*)} & \text{if } p \in (p_1^*, 1] \\ \frac{\lambda_0}{r+2\lambda_0} & \text{if } p \in [0, p_1^*] \end{cases}.$$

**Proof.** Proof is by a standard verification argument. Please refer to Appendix B for details, and Appendix A.1 for the ODE satisfied by the planner's value. ■

In the next subsection, we analyse the non-cooperative game between the firms.

## 3.2 Non-cooperative game

We restrict ourselves to Markov perfect equilibria, with the firms' common belief as the state variable. A Markov strategy for player $i$ ($i = 1, 2$) is defined as a left continuous function $k_i : [0, 1] \to \{0, 1\}$, $p \mapsto k_i(p)$. Let $v_i$ be the value function of player $i$. Given $k_j$ ($j \neq i$), player $i$'s Bellman equation is

$$rv_i = \max_{k_i\{0,1\}} \lambda_0[1-v] + k_i\{\lambda_1 p(1 - v_i - v_i' \cdot (1-p)) - \lambda_0(1-v_i)\}$$
$$- (1-k_j)\lambda_0 v_i - k_j \lambda_1 p(v_i + (1-p)v_i'). \quad (1)$$

From this Bellman equation, we can derive the best responses of the firms, using the ODEs exhibited in Appendix A.1.

Suppose firm $j \neq i$ is adopting method $R$ at the belief $p \in (0, 1)$. By left-continuity, there is a left-neighbourhood of $p$ in which $j$ is adopting $R$. If $i$ best-responds to $j$ by adopting $R$ in some subset of this left-neighbourhood, its value function satisfies

$$v_i \geq \frac{2\lambda_0 - \lambda_1 p}{r + 2\lambda_0}$$

on this subset. If the inequality is strict, adopting $R$ is $i$'s unique best response. By the same token, if the other firm is adopting the method $S$ in some left-neighbourhood of $p$, then, if firm $i$ best-responds by adopting the method $R$, its value function satisfies

$$v_i \geq \frac{\lambda_0}{r+2\lambda_0};$$

if the inequality is strict, adopting $R$ is $i$'s unique best response.

These simple observations allow us to prove the following result, which shows that the unique MPE in this setting coincides with the planner's solution.

**Proposition 2** *If firms are homogeneous, the unique MPE is efficient.*

**Proof.** See Appendix C. ■

# 4 Asymmetric Firms

In this section, we analyse the situation of firms that differ in their abilities to achieve a success by a good innovative method, i.e., $\lambda_1 > \lambda_2 > \lambda_0$. We again first analyse the problem of a benevolent social planner who seeks to maximise the firms' aggregate[5] discounted payoffs.

## 4.1 Benchmark: Social Planner's Problem

We can restrict the social planner to Markov strategies $k^t = (k_1^t, k_2^t)$ with the posterior belief $p_t$ as the state variable, where we write $k_i^t = 1(0)$ $(i = 1,2)$ if the planner assigns firm $i$ to method $R$ ($S$). The value function of the planner $v(p)$ satisfies

$$rv = \max_{k_i \in \{0,1\}(i=1,2)} 2\lambda_0(1-v) + k_1\left\{\lambda_1 p[1-v-v' \cdot (1-p)] - \lambda_0[1-v]\right\}$$
$$+ k_2\left\{\lambda_2 p[1-v-v'(1-p)] - \lambda_0[1-v]\right\} \tag{2}$$

The expression $2\lambda_0(1-v)$ is the expected flow payoff the planner can guarantee himself by using the method $S$. On the other hand, $\lambda_i p[1-v-v'(1-p)] - \lambda_0(1-v)$

[5]With asymmetric firms, it is convenient to do the analysis with aggregate, rather than average, payoffs.

reflects the premium the planner gets by assigning firm $i$ to method $R$. By linearity, it would be without loss for the planner to choose $\{k_i(p_t)\}_{t\geq 0}$ $(i=1,2)$ with $k_i(p_t)\in\{0,1\}$, even if firms' efforts were divisible. The following proposition describes the planner's solution. It shows that efficiency requires the planner to choose the myopically optimal method for firm 1, while assigning firm 2 to method $S$ also at some beliefs that are higher than its myopically optimal threshold. To state the theorem, we use the strictly decreasing and strictly convex functions $\mu(p)=(1-p)(\frac{1-p}{p})^{\frac{r}{\lambda_1+\lambda_2}}$ and $\mu_1(p)=(1-p)(\frac{1-p}{p})^{\frac{r+\lambda_0}{\lambda_1}}$.

**Proposition 3** *The planner's optimal solution is characterised by thresholds $p_1^*=\frac{\lambda_0}{\lambda_1}$ and $p_2^*\in(p_1^*,1)$, such that, for $p\in(p_2^*,1]$ ($p\in(0,p_1^*]$), both firms are assigned to method R (S). For $p\in(p_1^*,p_2^*]$, firm 1 is assigned to method R and firm 2 is assigned to method S. The planner's value function is given by*

$$v(p)=\begin{cases}\frac{\lambda_1+\lambda_2}{r+\lambda_1+\lambda_2}p+C_{rr}\mu(p) & \equiv v_{rr}(p) \quad \text{if } p\in(p_2^*,1],\\ \frac{\lambda_0}{r+\lambda_0}+\frac{r\lambda_1}{(r+\lambda_0)(r+\lambda_0+\lambda_1)}p+C_{rs}\mu_1(p) & \equiv v_{rs}(p) \quad \text{if } p\in(p_1^*,p_2^*],\\ \frac{2\lambda_0}{r+2\lambda_0} & \quad\quad\quad\quad\;\, \text{if } p\in[0,p_1^*].\end{cases} \tag{3}$$

*where $p_2^*\in(\frac{\lambda_0}{\lambda_2},1)$ satisfies*

$$v_{rr}(p_2^*)=v_{rs}(p_2^*)=\frac{\lambda_0(\lambda_1+\lambda_2)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}.$$

*$C_{rs}$ and $C_{rr}$ are constants of integration with $C_{rs}=\frac{r\lambda_0(\lambda_1-\lambda_0)}{(r+\lambda_0)(r+2\lambda_0)(r+\lambda_0+\lambda_1)u(p_1^*)}>0$, and $C_{rr}>0$ is determined from $v_{rr}(p_2^*)=v_{rs}(p_2^*)$.*

**Proof.**

Proof is by a standard verification argument. Please refer to Appendix (D) for details, and Appendix A.2 for the ODEs satisfied by the planner's value function.

■

The planner's value function $v(p)$ is of class $C^1$, (strictly) increasing and (strictly) convex (on $(p_1^*,1)$). At the optimum, firm 2 switches from method $R$ to method $S$ as soon as the belief drops below the threshold $p_2^*>\frac{\lambda_0}{\lambda_2}$. By contrast, if firm 2 was the only firm around, then it would have optimally switched to method $S$ at the belief $\frac{\lambda_0}{\lambda_2}$. In the presence of firm 1, however, firm 2 optimally switches at a belief higher than its myopic threshold, while firm 1 optimally switches to method $S$ at its myopic threshold $\frac{\lambda_0}{\lambda_1}$. Thus the planner has firm 2 switch its action at a belief where the expected arrival rate on the innovative method is higher than that of the safe method.

This a priori surprising information aversion by the planner can be intuitively explained as follows. Since the game ends after the first breakthrough, there is no learning motive and hence, in the planner's solution no firm will be made to use method $R$ for beliefs less than its myopic cutoff. This implies firm 1 is the last firm to switch to method $S$ at its myopic belief $p_1^* = \frac{\lambda_0}{\lambda_1}$. Since firm 1 is more productive than firm 2, the planner would gain if he could contemporaneously substitute firm 1's experimentation for firm 2's. While such a contemporaneous substitution is not feasible, it is however indeed possible for the planner to substitute *future* experimentation by firm 1 for *current* experimentation by firm 2. For any belief strictly greater than $\frac{\lambda_0}{\lambda_1}$, while more future experimentation by firm 1 leads to an expected positive gain, the planner incurs an expected loss by giving up current experimentation by firm 2. At the myopic belief of firm 2 ($\frac{\lambda_0}{\lambda_2}$), this expected loss is equal to zero. This explains why the cutoff $p_2^*$ is strictly greater than $\frac{\lambda_0}{\lambda_2}$. Formally this can be understood as follows. At any belief, the expected positive gain from making 2 use $R$ is $(\lambda_2 p - \lambda_0)(1-v)$, and the expected loss from the environment becoming more pessimistic following hapless experimentation by 2 is $-\lambda_2 p(1-p)v'$. Since $v$ is strictly convex and increasing in $p$ for $p \in (p_1^*, 1)$, at $p = \frac{\lambda_0}{\lambda_2}$, we have $v'(p) > 0$. This implies that, at $p = \frac{\lambda_0}{\lambda_2}$, the expected gain $(\lambda_2 p - \lambda_0)(1-v) = 0$ is outweighed by the cost $-\lambda_2 p(1-p)v' < 0$.

## 4.2 Non-cooperative game

In this section, we will discuss and analyse the non-cooperative game between the firms. Our solution concept is Markov perfect equilibrium. Given $k_j$ ($j = 1, 2$), if $v_i$

$(i = 1,2; i \neq j)$ is the payoff of firm $i$ in equilibrium, then we have

$$v_i = \max_{k_i \in \{0,1\}} \left\{ \lambda_0 (1-k_i)\, dt + k_i \lambda_i p\, dt \right.$$

$$\left. +(1-r\,dt)[1-\lambda_0(1-k_i)\,dt-(1-k_j)\lambda_0\,dt-(k_i\lambda_i+k_j\lambda_j)p\,dt][v_i-(k_i\lambda_i+k_j\lambda_j)p(1-p)v_i^{'}dt] \right\}$$

$$\Rightarrow rv_i = \lambda_0(1-v_i) + \max_{k_i \in \{0,1\}} k_i\{\lambda_i p[1-v_i-v_i^{'}(1-p)] - \lambda_0(1-v_i)\}$$

$$-(1-k_j)\lambda_0 v_i - k_j\{\lambda_j p[v_i + v_i^{'}(1-p)]\} \tag{4}$$

Each firm $i$ $(i = 1,2)$ can guarantee itself an expected flow payoff of $\lambda_0(1-v_i)$ by using the traditional method ($S$). The term $\{\lambda_i p[1-v_i-v_i^{'}(1-p)] - \lambda_0(1-v_i)\}$ captures the premium firm $i$ receives by using the innovative method. The expressions $-(1-k_j)\lambda_0 v_i$ and $-k_j\{\lambda_j p[v_i + v_i^{'}(1-p)]\}$ account for the negative payoff externalities.

**Best Responses:**

Suppose $k_j = 0$ $(j \in \{1,2\})$ in an open neighborhood of $p$. From (4), we can see that using method $R$ in a neighborhood of $p$ is optimal for firm $i$ $(i \in \{1,2\}; i \neq j)$ if and only if

$$v_i \geq \frac{\lambda_0}{r+2\lambda_0}$$

is satisfied in that neighborhood.

Next, suppose $k_j = 1$ in an open neighborhood of $p$. From (4), we can infer that choosing $R$ is optimal for firm $i$ in a neighborhood of $p$ if and only if

$$v_i \geq \frac{\lambda_0(\lambda_1+\lambda_2) - \lambda_1\lambda_2 p}{r\lambda_i + \lambda_0(\lambda_1+\lambda_2)} \tag{5}$$

is satisfied in that neighborhood.

**Proposition 4** *There exists a unique Markov perfect equilibrium. In this equilibrium, firm* 1 *uses the innovative method (R) in the belief region* $(\frac{\lambda_0}{\lambda_1}, 1]$*, and the safe method (S) otherwise. Firm* 2 *uses the innovative method on* $(\hat{p}_2, 1[$ *and the safe* *method (S) otherwise, where the cutoff $\hat{p}_2$ satisfies*

$$\frac{\lambda_0}{\lambda_1} < \hat{p}_2 < \frac{\lambda_0}{\lambda_2}.$$

*The firms' equilibrium payoffs are given by*

$$v_1(p) = \begin{cases} v_1^{rr}(p) = \frac{\lambda_1}{r+\lambda_1+\lambda_2}p + C_1^{rr}\mu(p) & : \text{ if } p \in (\hat{p}_2, 1] \\ v_1^{rs}(p) = \frac{\lambda_1}{r+\lambda_0+\lambda_1}p + C_1^{rs}\mu_1(p) & : \text{ if } p \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2] \\ \frac{\lambda_0}{r+2\lambda_0} & : \text{ if } p \in (0, \frac{\lambda_0}{\lambda_1}], \end{cases} \tag{6}$$

*and*

$$v_2(p) = \begin{cases} v_2^{rr}(p) = \frac{\lambda_2}{r+\lambda_1+\lambda_2}p + C_2^{rr}\mu(p) & : \text{ if } p \in (\hat{p}_2, 1], \\ v_2^{rs}(p) = \frac{\lambda_0}{r+\lambda_0}(1 - \frac{\lambda_1}{r+\lambda_0+\lambda_1}p) + C_2^{rs}\mu_1(p) & : \text{ if } p \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2] \\ \frac{\lambda_0}{r+2\lambda_0} & : \text{ if } p \in (0, \frac{\lambda_0}{\lambda_1}], \end{cases} \tag{7}$$

*respectively. The constants of integration are determined by value matching: $C_1^{rs} > 0$ is given by $v_1^{rs}(p_1^*) = \frac{\lambda_0}{r+2\lambda_0}$ and $C_2^{rs} < 0$ by $v_2^{rs}(p_1^*) = \frac{\lambda_0}{r+2\lambda_0}$. The threshold $\hat{p}_2$ is defined implicitly by $v_2^{rs}(\hat{p}_2) = \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2\hat{p}_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. The constants of integration $C_1^{rr} > 0$ and $C_2^{rr} > 0$ are determined by $v_1^{rr}(\hat{p}_2) = v_1^{rs}(\hat{p}_2)$, and $v_2^{rr}(\hat{p}_2) = v_2^{rs}(\hat{p}_2)$, respectively. The function $v_2$ is smooth, while $v_1$ is smooth everywhere except at $p = \hat{p}_2$.*

**Proof.** Existence follows from standard verification arguments (please refer to Appendix (E.1) for details). Uniqueness follows from the Bellman equation (4) and the relevant ODE's (Appendix (A.2)). (Please see Appendix (E.2) for a detailed proof.) ■

Firm 1's value function is increasing and convex throughout. Firm 2's value function, by contrast, is concave when only firm 1 uses method *R*; it is convex when both firms use it. It is constant when both firms use method *S*, then becomes decreasing in the range where only firm 1 uses method *R*. It has an inflection point at $\hat{p}_2$, where firm 2 switches methods, and eventually becomes increasing as firms become very optimistic about method *R*.

Note that, in the unique Markov perfect equilibrium, both firms are using a cutoff strategy, that is they use the innovative method if and only if the likelihood of it being good is above a threshold.

As shown in appendix (E.1), the slope of the payoff function of firm 2 at $p = \hat{p}_2$ is strictly negative. From the Bellman equation (4) we can infer that

$$\lambda_2 \hat{p}_2 [1 - v_2 - v_2^{'}(1 - \hat{p}_2)] = \lambda_0 (1 - v_2)$$

which implies that $\hat{p}_2 < \frac{\lambda_0}{\lambda_2}$, i.e., firm 2 experiments *more* than if it were acting myopically. By contrast, we have seen that $p_2^* > \frac{\lambda_0}{\lambda_2}$, i.e., a utilitarian planner would prefer firm 2 to experiment *less* than if it were myopic. Thus, the unique Markov perfect equilibrium involves inefficient duplicative use of method $R$ on $(\hat{p}_2, p_2^*)$. This result can be intuitively explained as follows. Because of the winner-takes-all structure, both firms want to be the first inventor. At the belief $p = \frac{\lambda_0}{\lambda_2}$, the myopic payoff to firm 2 is the same for both methods. However, by using method $R$, firm 2 is producing additional information, implying that, if there is no breakthrough, firms become more pessimistic about the innovative method. In equilibrium, though, firm 1 uses the method $R$ until the belief reaches $p_1^* = \frac{\lambda_0}{\lambda_1}$. Thus, as the belief decreases due to firm 2's unsuccessful use of method $R$, the time firm 1 spends using $R$ is reduced. Based on the current belief $\frac{\lambda_0}{\lambda_2}$, this reduces the chances of a breakthrough by firm 1, thereby increasing the future expected payoff of firm 2. Thus, firm 2 has an incentive to use method $R$ for some beliefs below $\frac{\lambda_0}{\lambda_2}$.

# 5 Comparative Statics

## 5.1 Degree of Heterogeneity and Inefficiency

For given values of the parameters $\lambda_0$, $\lambda_1$ and $r$, we define $e(\lambda_2) = p_2^* - \hat{p}_2$ as a measure of the amount of inefficient duplication as a function of firm heterogeneity. Let $e^* = \sup_{\lambda_2 \in [\lambda_0, \lambda_1]} e(\lambda_2)$. In the following proposition, we shall show that inefficient duplication is worst for intermediate degrees of heterogeneity, i.e., $e^*$ is attained at intermediate values of $\lambda_2$. To state the proposition, we define, for each $\varepsilon > 0$, the set $\Lambda_2(\varepsilon) = \{\lambda_2 \in [\lambda_0, \lambda_1] : \text{s.t. } e(\lambda_2) < \varepsilon\}$, i.e., the set of productivities of the weak firm such that our measure of inefficient duplication is $\varepsilon$-close to its supremum.

**Proposition 5** *There exists $\bar{\varepsilon} > 0$ such that there exist $\lambda_2^1 \in (\lambda_0, \lambda_1)$ and $\lambda_2^2 \in (\lambda_2^1, \lambda_1)$ such that $\Lambda_2(\varepsilon) \subset (\lambda_2^1, \lambda_2^2)$ for all $\varepsilon \in (0, \bar{\varepsilon})$.*

**Proof.** For $\lambda_2 = \lambda_0$, in both the planner's problem and the non-cooperative game, firm 2 chooses method $S$ for any $p < 1$, implying $p_2^* = \hat{p}_2 = 1$, and hence $e(\lambda_0) = 0$. By the same token, when $\lambda_2 = \lambda_1$, Proposition 2 implies $p_2^* = \hat{p}_2 = \frac{\lambda_0}{\lambda_1}$, and thus $e(\lambda_1) = 0$. For any $\lambda_2 \in (\lambda_0, \lambda_1)$, by contrast, $e(\lambda_2) > 0$ by Proposition 4. The claim thus follows by continuity of $e(\cdot)$. ■

Thus, the severity of inefficient duplication is maximal when the degree of heterogeneity among the firms is intermediate. For high degrees of heterogeneity, the lowest belief at which firm 2 chooses the innovative method is close to 1 in both the planner's problem and the non-cooperative game, and hence the range of beliefs over which inefficient duplication occurs is low. Since with homogeneous players, the unique equilibrium is efficient, continuity implies a small range of inefficient duplication when the productivities of the firms are not too different from each other. Fig 1 depicts $e(\lambda_2)$ for $\lambda_0 = 1$, $\lambda_1 = 5$, and $r = 0.4$.

## 5.2 Promoting early completion

The following proposition shows that the inefficient duplication of efforts is attenuated if actors are very impatient. We can derive an important policy prescription from this result. A funding agency awarding research grants to competing researchers, for instance, could mitigate the problem of inefficient duplicative search by providing incentives for early completion. In this subsection, we fix $\lambda_0$, $\lambda_1$ and $\lambda_2$, and write $p_2^*$ and $\hat{p}_2$ as functions of the discount rate $r$.

**Proposition 6** *Fix $r_0 > 0$. There exists $\bar{r} \in (r_0, \infty)$ such that, for all $r > \bar{r}$, we have $p_2^*(r) < p_2^*(r_0)$ and $\hat{p}_2(r) > \hat{p}_2(r_0)$.*

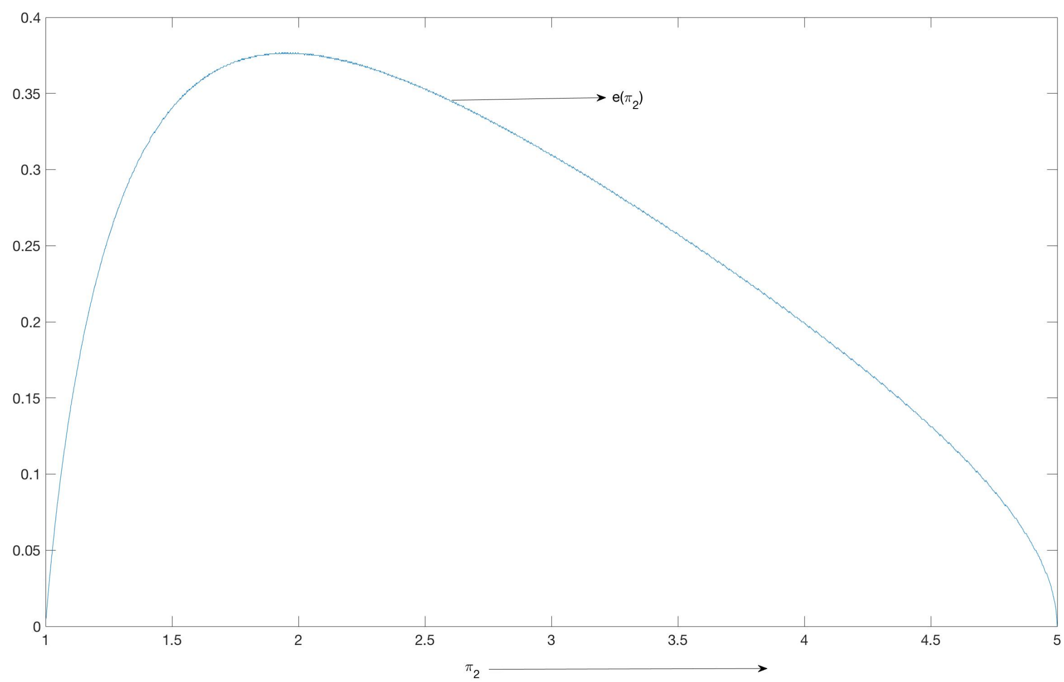


Figure 1: Size of the range of inefficient duplication as a function of $\lambda_2$

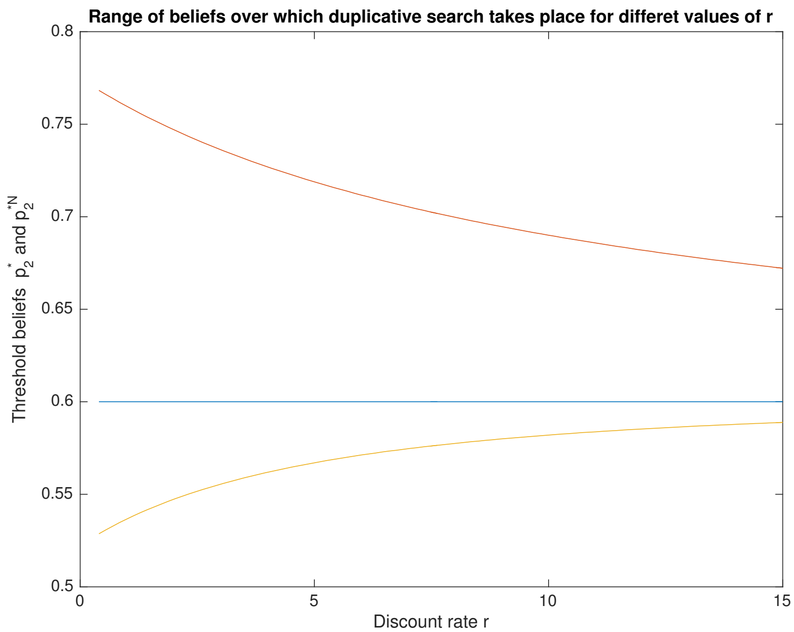


Figure 2: Range of beliefs over which duplicative search takes place.

**Proof.** Please see Appendix F. ■

There is duplicative search in equilibrium over the range $(\hat{p}_2, p_2^*)$. Thus we have shown that, when $r$ increases sufficiently, the range of beliefs over which there is duplicative search shrinks. The intuition is as follows. Equilibrium inefficiency arises here from the fact that firm 2 cares only about its private expected future payoff and consequently disregards the negative externality it inflicts by excessively pursuing the innovative avenue. As players become myopic, the importance of expected future payoffs decreases and hence the extent of the excessive search is reduced.

Figure 2 shows how $p_2^*$ and $\hat{p}_2$ change with $r$, for parameter values $\lambda_0 = 3$, $\lambda_1 = 8$, and $\lambda_2 = 5$. The horizontal straight line represents the myopic belief $\frac{\lambda_0}{\lambda_2}$. The curve above it depicts the values of $p_2^*$ while the curve below it depicts the values of $\hat{p}_2$ as a function of $r$.

# 6 Conclusion

We have shown that in a patent race model with dynamic learning and optimal readjustment of project selection, the combination of payoff externalities and heterogeneous players gives rise to equilibrium inefficiency in the form of too much experimentation. The problem is particularly prevalent for intermediate levels of competitors' heterogeneity. Our analysis would suggest that, in order to avoid inefficient duplication, research agencies should especially encourage early successes.

In contrast to models of purely informational externalities such as Keller, Rady and Cripps (2005), our Markov perfect equilibrium is unique. It is furthermore in cutoff strategies, while there does not exist an equilibrium in cutoff strategies in Keller, Rady and Cripps (2005). Moreover, inefficiency in our setting arises because of excessive information production, while all equilibria in Keller, Rady and Cripps (2005) are inefficient because players experiment too little in equilibrium.

In our model, research abilities, and hence the degree of player heterogeneity, were exogenously given. It would be interesting to investigate a setting in which players' abilities grew over time as a function of past research efforts (learning by doing). Furthermore, the decision to take out a patent, and thus to make one's findings public, is often a strategic decision, conceivably impacting firms' choices of research avenues. We commend these questions to future research.

# References

AKCIGIT, U. and LIU, Q. (2011): "The Role of Information in Competitive Experimentation," *Journal of the European Economic Association*, 14(4), 828-870.

BESANKO, D. and WU, J. (2013): "The Impact of Market Structure and Learning on the Tradeoff Between R & D Competition and Cooperation," *The Journal of Industrial Economics*, LXI, 166-201.

BHATTACHARYA, S. and MOOKERJEE D. (1986): "Portfolio Choice in Research and Development," *Rand Journal of Economics*, 17, 594-605.

BOLTON, P. and HARRIS, C. (1999): “Strategic Experimentation,” *Econometrica*, 67, 349-374.

CHATTERJEE, K. and EVANS, R. (2004): “Rivals’ Search for Buried Treasure: Competition and Duplication in R & D,” *Rand Journal of Economics*, 35, 160-183.

DAS, K., KLEIN, N. and SCHMID, K. (2019): “Strategic Experimentation with Asymmetric Players,” *Economic Theory*, Open Access, available at https://link.springer.com/article/10.1007/s00199-019-01193-9.

DASGUPTA, P. and MASKIN, E. (1987): “The Simple Economics of Research Portfolios,” *The Economic Journal*, 97(387), 581-595.

DONG, H., YUEDE, C.M., COUGHLAN C.A., MURPHY K.M. and CSERNANSKY J G. (2009): “Effects of Donepezil on Amyloid-$\beta$ and Synapse Density in the Tg2576 Mouse Model of Alzheimer’s Disease,” *Brain Research*, 1303, 169-178.

EDERER, F. (2013): “Incentives for Parallel Innovation,” mimeo, Yale University.

FERSHTMAN, C. and RUBINSTEIN, A. (1997): “A Simple Model of Equilibrium in Search Procedures,” *Journal of Economic Theory*, 72, 432-441.

HOPENHAYN, H. and SQUINTANI, F. (2016): “The Direction of Innovation,” mimeo, UCLA and University of Warwick.

HÖRNER, J. and SKRZYPACZ, A. (2016): “Learning, Experimentation and Information Design,” working paper, Stanford University.

KELLER, G. and S. RADY (2010): “Strategic Experimentation with Poisson Bandits,” *Theoretical Economics*, 5, 275–311.

KELLER, G., RADY, S. and CRIPPS,M. (2005): “Strategic Experimentation with Exponential Bandits,” *Econometrica*, 73, 39–68.

KLEIN, N. (2013): "Strategic Learning in Teams," *Games and Economic Behavior*, 82, 636–657.

KLEIN, N. (2016): "The Importance of Being Honest," *Theoretical Economics*, 11, 773-811.

KLEIN, N. and S. RADY (2011): "Negatively Correlated Bandits," *Review of Economic Studies*, 78(2), 693–732.

LEE, T. and WILDE, L. (1980): "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economics*, 94, 429-436.

LETINA, I. (2016): "The Road not Taken: Competition and the R & D Portfolio," *Rand Journal of Economics*, 47(2), 433-460.

MANSO, G. (2011): "Motivating Innovation," *The Journal of Finance*, LXVI 5, 1823-1860.

MATROS, A. and SMIRNOV, V. (2016): "Duplicative Search," *Games and Economic Behavior*, 99, 1-22.

MOSS, D. (2018): "Cholinesterase Inhibitors–The Future of Alzheimer's Therapy?," *SciTech Europa Quarterly*, 27, 208-209.

PAMMOLLI, F., MAGAZZINI, L. and RICCABONI, M. (2011): "The Productivity Crisis in Pharmaceutical R & D," *Nature Reviews Drug Discovery*, 10(6), 428.

PRESMAN, E.L. (1990): "Poisson Version of the Two-Armed Bandit Problem with Discounting," *Theory of Probability and its Applications*, 35, 307–317.

REINGANUM, J. (1982): "A Dynamic Game of R & D Patent Protection and Competitive Behavior," *Econometrica*, 50, 671-688.

SCHERER, F.M. (2000): *International High-Technology Competition*. Harvard University Press, Cambridge, Mass.

THOMAS, C. (2017): “Experimentation with Congestion,” mimeo, University of Texas, Austin.

# APPENDIX

# A Ordinary Differential Equations

We define the following decreasing and convex functions:

$$\mu_i(p) = (1-p)\left(\frac{1-p}{p}\right)^{\frac{r+\lambda_0}{\lambda_i}};$$

$$\mu(p) = (1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda_1+\lambda_2}}.$$

Throughout this section, we write $C$ for a constant of integration, which is determined from the specific boundary condition. We furthermore write $i$ and $j$ for the two firms, i.e, $\{i,j\} = \{1,2\}$.

## A.1 ODEs in the game with symmetric firms

**Planner's problem:**

If $k = 0$ is chosen at belief $p$, the planner's payoff satisfies $v(p) = \frac{\lambda_0}{r+2\lambda_0}$. If the planner chooses $k = 2$ on an open set of beliefs, his payoff function satisfies the ODE

$$2\lambda_1 p(1-p)v' + (r+2\lambda_1 p)v = \lambda_1 p. \tag{8}$$

This is solved by

$$v(p) = \frac{\lambda_1 p}{r+2\lambda_1} + C\mu(p). \tag{9}$$

**The non-cooperative game:**

Suppose firm $i$ adopts method $R$.

If firm $j$ adopts method $R$ as well, either player's value function satisfies the ODE (8), which, as we have seen above, admits the solution (9). If firm $j$ adopts the method $S$, by contrast, firm $i$'s value function satisfies

$$\lambda_i p(1-p)v_i^{'} + (r+\lambda_0+\lambda_i p)v_i = \lambda_i p. \tag{10}$$

This is solved by

$$v_i(p) = \frac{\lambda_i}{r+\lambda_0+\lambda_i} p + C\mu_i(p). \tag{11}$$

If now $i$ adopts $S$ and $j$ adopts $R$, then $i$'s value function satisfies

$$\lambda_j p(1-p)v_i^{'} + (r+\lambda_0+\lambda_j p)v_i = \lambda_0. \tag{12}$$

This is solved by

$$v_i(p) = \frac{\lambda_0}{r+\lambda_0}(1 - \frac{\lambda_j}{r+\lambda_0+\lambda_j} p) + C\mu_j(p). \tag{13}$$

Finally, if both firms adopt $S$, either firm's value function satisfies

$$v(p) = \frac{\lambda_0}{r+2\lambda_0}.$$

## A.2 ODEs in the game with asymmetric firms

**Planner's problem:**

If $k_1 = k_2 = 0$ at belief $p$, the planner's payoff is $v(p) = \frac{2\lambda_0}{r+2\lambda_0}$.

If the planner chooses $k_1 = 1$ and $k_2 = 0$ on an open set of beliefs, his payoff function satisfies the ODE

$$\lambda_1 p(1-p)v^{'} + (r+\lambda_0+\lambda_1 p)v = \lambda_0 + \lambda_1 p. \tag{14}$$

This is solved by

$$v(p) = \frac{\lambda_0}{r+\lambda_0} + \frac{r\lambda_1}{(r+\lambda_0)(r+\lambda_0+\lambda_1)} p + C\mu_1(p). \tag{15}$$

If the planner chooses $k_1 = k_2 = 1$ on an open set of beliefs, his payoff function satisfies the ODE

$$(\lambda_1 + \lambda_2)p(1-p)v^{'} + (r + (\lambda_1 + \lambda_2)p)v = (\lambda_1 + \lambda_2)p. \tag{16}$$

This is solved by

$$v(p) = \frac{\lambda_1 + \lambda_2}{r + \lambda_1 + \lambda_2}p + C\mu(p). \tag{17}$$

**Non-cooperative game**

Suppose both firms adopt method $S$. Inserting $k_1 = k_2 = 0$ in (4), we can see that both players' payoff is given by the constant

$$\frac{\lambda_0}{r + 2\lambda_0}. \tag{18}$$

Suppose firm $i$ adopts method $R$ and $j$ adopts method $S$. Inserting $k_i = 1$ and $k_j = 0$ in (4), we can infer that the payoff function of firm $i$ satisfies the ODE

$$\lambda_i p(1-p)v_i^{'} + (r + \lambda_0 + \lambda_i p)v_i = \lambda_i p. \tag{19}$$

The solution to the above differential equation is

$$v_i^{rs}(p) = \frac{\lambda_i}{r + \lambda_0 + \lambda_i}p + C\mu_i(p). \tag{20}$$

Firm $j$'s payoff satisfies

$$\lambda_i p(1-p)v_j^{'} + (r + \lambda_0 + \lambda_i p)v_j = \lambda_0. \tag{21}$$

The solution to the above differential equation is

$$v_j^{rs}(p) = \frac{\lambda_0}{r + \lambda_0}\left(1 - \frac{\lambda_i}{r + \lambda_0 + \lambda_i}p\right) + C\mu_i(p). \tag{22}$$

Finally, consider the situation where both firms adopt method $R$. Inserting $k_1 =$

$k_2 = 1$ in (4), we can infer that the payoff function of either firm $i$ satisfies the ODE

$$(\lambda_1 + \lambda_2)p(1-p)v_i^{'} + (r + (\lambda_1 + \lambda_2)p)v_i = p\lambda_i. \tag{23}$$

The solution to the above differential equation is

$$v_i^{rr}(p) = \frac{\lambda_i}{r + \lambda_1 + \lambda_2} p + C\mu(p). \tag{24}$$

# B Proof of Proposition 1

The payoff function associated with the policy $k^*$ is $v$. Since $\frac{\lambda_0}{r+2\lambda_0} - \frac{\lambda_1}{r+2\lambda_1} p_1^* > 0$, we know that for $p \in (p_1^*, 1)$, $v$ is strictly convex. Since $v$ satisfies the value matching condition at $p = p_1^*$, direct computation shows that $v^{'}(p_1^*) = 0$. Hence, $v$ is of class $C^1$ and strictly increasing for $p \in (p_1^*, 1)$. From the ODE (8), we know that $\lambda_1 p[\frac{1}{2} - v - v^{'}(1-p)] = \frac{r}{2}v$. At $p = p_1^*$, $v = \frac{\lambda_0}{r+2\lambda_0}$. This implies $rv = \lambda_0(1-2v)$. Since $v$ is strictly increasing for $p > p_1^*$, for all $p \in (p_1^*, 1)$ we have $rv > \lambda_0(1-2v) \Rightarrow \lambda_1 p[1 - 2v - 2v^{'}(1-p)] > \lambda_0(1-2v)$. Thus, choosing $k = 2$ solves the Bellman equation. On the other hand, since $v^{'} = 0$ for $p \leq p_1^*$, we have $\lambda_1 p[1 - 2v - 2v^{'}(1-p)] \leq \lambda_0(1-2v)$ for $p \in (0, p_1^*]$. Hence, choosing $k = 0$ satisfies the Bellman equation. This shows that the payoff function associated with the proposed policy satisfies the Bellman equation, and hence constitutes the planner's value function.

# C Proof of Proposition 2

We will show that given firm $j$ ($j = 1, 2$) adopts the method $R$ for $p > p_1^*$ and $S$ for $p \leq p_1^*$, this strategy also constitutes the best response of firm $i$. Consider $p \leq p_1^*$. In this range, we have $v_i = \frac{\lambda_0}{r+2\lambda_0}$. Given firm $j$'s strategy, $i$ has no incentive to deviate as $\lambda_1 p[1 - \frac{\lambda_0}{r+2\lambda_0}] < \lambda_0[1 - \frac{\lambda_0}{r+2\lambda_0}]$ for $p < p_1^*$. Next, consider the range of beliefs $(p_1^*, 1)$. From the closed-form solution of $v_i$ (see equation 9 in Appendix A.1 ) we can see that $v_i$ is strictly increasing and convex as $[\frac{\lambda_0}{r+2\lambda_0} - \frac{\lambda_1}{r+2\lambda_1} p_1^*] > 0$. At $p = p_1^*$, $v_i = \frac{2\lambda_0 - \lambda_1 p}{r+2\lambda_0}$. Since $\frac{2\lambda_0 - \lambda_1 p}{r+2\lambda_0}$ is strictly decreasing in $p$, for all $p > p_1^*$ we have $v_i > \frac{2\lambda_0 - \lambda_1 p}{r+2\lambda_0}$.

To show uniqueness, consider again the range $p \leq p_1^*$ and suppose that a firm adopts the method $R$ for a range of beliefs $(p_l, p_h)$ such that $p_l < p_h \leq p_1^*$. Let $\hat{p} < p_1^*$ be the infimum of such beliefs $p_l$. Then, $v_j(\hat{p}) = v_i(\hat{p}) = \frac{\lambda_0}{r+2\lambda_0}$. Assume without loss of generality that firm $i$ adopts method $R$ in some right-neighbourhood of $\hat{p}$. By the ODEs (8) and (10), it follows immediately from $\hat{p} < p_1^*$ that $v_i < \frac{\lambda_0}{r+2\lambda_0} < \frac{2\lambda_0 - \lambda_1 p}{r+2\lambda_0}$ to the immediate right of $\hat{p}$, implying $i$ has a profitable deviation in a right-neighbourhood of $\hat{p}$.

Now, consider the range $(p_1^*, 1]$. We shall first show that there cannot be a $\check{p} \in (p_1^*, 1]$ such that $(k_i, k_j)(\check{p}) = (0,0)$ in any equilibrium. Indeed, suppose to the contrary that this was the case. Then, $v_i(\check{p}) = v_j(\check{p}) = \frac{\lambda_0}{r+2\lambda_0}$. By left-continuity of strategies, there exists some left-neighbourhood $\mathscr{N}$ of $\check{p}$ such that $v_i = v_j = \frac{\lambda_0}{r+2\lambda_0}$ and $v_i' = v_j' = 0$ in this neighbourhood. The Bellman equation (1) now implies that either player has a profitable deviation on $\mathscr{N} \cap (p_1^*, \check{p})$. Next, we shall show that, in any equilibrium, $(k_i, k_j) = (1,1)$ prevails in some right-neighbourhood of $p_1^*$. Suppose to the contrary that (without loss) $(k_i, k_j) = (1,0)$ prevails in some right-neighbourhood of $p_1^*$ in some equilibrium. Then, by (12), $v_j'(p_1^*) = 0$, implying $v_j > \frac{2\lambda_0 - \lambda_1 p}{r+2\lambda_0}$ to the immediate right of $p_1^*$. Thus, $j$ has a profitable deviation to the immediate right of $p_1^*$. Now, suppose there is an equilibrium in which it is not the case that $(k_i, k_j) = (1,1)$ prevails everywhere on $(p_1^*, 1]$. Then, there exists some $\tilde{p} \in (p_1^*, 1]$ and a firm $j$ such that $v_j(\tilde{p}) = \frac{\lambda_0}{r+2\lambda_0}$ and $v_j'(\tilde{p}-) \leq 0$. (8) and (10) imply that we must have $(k_i, k_j)(\tilde{p}) = (0,1)$. Yet, by (12), $\tilde{p} > p_1^*$ implies $v_j'(\tilde{p}-) < 0$, so that $j$ has a profitable deviation to the immediate left of $\tilde{p}$.

# D Proof of Proposition 3

The policy $k^* = (k_1^*, k_2^*)$ implies the payoff function $v$ (given by (3)). As $C_{rs} > 0$, $v_{rs}(p_1^*) = \frac{2\lambda_0}{r+2\lambda_0}$ and $v_{rs}'(p_1^*) = 0$, $v|_{(0,p_2^*)}$ is $C^1$, (strictly) increasing and (strictly) convex (on $(p_1^*, p_2^*)$). By ODEs (14) and (16), we have that $v_{rs}'(p_2^*) = v_{rr}'(p_2^*)$. We shall now show that this smooth pasting at $p_2^*$ implies that $C_{rr} > 0$. Indeed, assume to the contrary that $C_{rr} \leq 0$. As $\mu_h' < 0$ and $p_2^* < 1$, this implies $v_{rr}'(p_2^*) > \frac{\lambda_1 + \lambda_2}{r+\lambda_1+\lambda_2}$. Yet, as $C_{rs} > 0$ and $\mu_1' < 0$, we have that $v_{rs}'(p_2^*) < \frac{r\lambda_1}{(r+\lambda_0)(r+\lambda_0+\lambda_1)} < \frac{\lambda_1+\lambda_2}{r+\lambda_1+\lambda_2}$, a contradiction. Thus, $C_{rr} > 0$, and the payoff function $v$ is $C^1$, (strictly) increasing and (strictly) convex (on $(p_1^*, 1)$).

On $(0, p_1^*)$, $v = \frac{2\lambda_0}{r+2\lambda_0}$ and $v' = 0$, so that $\lambda_i p(1-v) - \lambda_0(1-v) < 0$, as $p < p_1^* = \frac{\lambda_0}{\lambda_1} < \frac{\lambda_0}{\lambda_2}$. Thus, $k_1^* = k_2^* = 0$ solves the Bellman equation (2) in this range.

For $p \in (p_1^*, p_2^*)$, (14) implies

$$\lambda_1 p[1 - v - v^{'}(1-p)] = (r + \lambda_0)v - \lambda_0$$

Since $v(p_1^*) = \frac{2\lambda_0}{r+2\lambda_0}$ and $v$ is strictly increasing on $(p_1^*, p_2^*)$, we have $\lambda_1 p[1 - v - v^{'}(1-p)] = (r+\lambda_0)v - \lambda_0 > \lambda_0(1-v)$ for this range of beliefs. Thus, $k_1^* = 1$ solves (2) for these beliefs. By the same token, (14) gives us

$$\lambda_2 p[1 - v - v^{'}(1-p)] = \frac{\lambda_2}{\lambda_1}[(r+\lambda_0)v - \lambda_0].$$

Since $v$ is strictly increasing on $(p_1^*, p_2^*)$ and $v(p_2^*) = v_{rs}(p_2^*) = v_{rr}(p_2^*) = \frac{\lambda_0(\lambda_1+\lambda_2)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$, we have that $v < \frac{\lambda_0(\lambda_1+\lambda_2)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ in this range, and hence

$$\lambda_2 p[1 - v - v^{'}(1-p)] = \frac{\lambda_2}{\lambda_1}[(r+\lambda_0)v - \lambda_0] < \lambda_0(1-v).$$

Hence, $k_2^* = 0$ solves (2) on $(p_1^*, p_2^*)$.

Now, let $p > p_2^*$. As $v$ is strictly increasing, $v(p) > \frac{\lambda_0(\lambda_1+\lambda_2)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} = v(p_2^*) > \frac{\lambda_0(\lambda_1+\lambda_2)}{r\lambda_1+\lambda_0(\lambda_1+\lambda_2)}$. By (16), we have

$$\lambda_2 p[1 - v - v^{'}(1-p)] = \frac{\lambda_2}{\lambda_1+\lambda_2} rv,$$

and hence $\lambda_i p[1 - v - v^{'}(1-p)] > \lambda_0(1-v)$ $(i = 1, 2)$. Thus, $k_1^* = k_2^* = 1$ solves (2) for $p > p_2^*$.

In conclusion, the payoff function $v$ is $C^1$, and solves the Bellman equation (2); it is thus the value function, and $k^* = (k_1^*, k_2^*)$ is the optimal policy.

It remains to show that $p_2^* > \frac{\lambda_0}{\lambda_2}$. From (2), we can infer that

$$\lambda_2 p_2^*[1 - v(p_2^*) - (1-p_2^*)v^{'}(p_2^*)] = \lambda_0[1 - v(p_2^*)]$$

Since $v^{'}(p_2^*) > 0$ and $v(p_2^*) < 1$, we have

$$\lambda_2 p_2^*[1 - v(p_2^*)] > \lambda_2 p_2^*[1 - v(p_2^*) - (1 - p_2^*)v^{'}(p_2^*)] = \lambda_0[1 - v(p_2^*)]$$

$$\Rightarrow p_2^* > \frac{\lambda_0}{\lambda_2}.$$

# E Proof of Proposition 4

## E.1 Verification Argument

The proposed policies imply a well-defined law of motion of the posterior belief, and lead to the payoff functions as stated in the theorem.

The constant of integration $C_1^{rs}$ is determined from $v_1^{rs}\left(\frac{\lambda_0}{\lambda_1}\right) = \frac{\lambda_0}{r+2\lambda_0}$, which immediately implies $C_1^{rs} > 0$, as $\lambda_1 > \lambda_0$. Since $v_1^{rs}(\hat{p}_2) = v_1^{rr}(\hat{p}_2)$, it follows from $\lambda_2 > \lambda_0$ that $C_1^{rr} > 0$ as well. Direct computation shows $v_1^{rs'}(\frac{\lambda_0}{\lambda_1}+) = 0$. Furthermore, the ODEs (19) and (23) imply that $v_1'(\hat{p}_2+) > \frac{\lambda_1 v_1'(\hat{p}_2-)}{\lambda_1+\lambda_2} > 0$, for $\hat{p}_2 < \frac{\lambda_0}{\lambda_2}$. (This inequality will be shown below). We can thus conclude that $v_1$ is continuously differentiable anywhere except at $\hat{p}_2$ and (strictly) increasing and convex (in the range $[\frac{\lambda_0}{\lambda_1}, 1]$).

By the same token, the constant of integration $C_2^{rs}$ is determined from $v_2^{rs}\left(\frac{\lambda_0}{\lambda_1}\right) = \frac{\lambda_0}{r+2\lambda_0}$. Direct calculation shows that this implies $C_2^{rs} < 0$ and $v_2^{rs'}(\frac{\lambda_0}{\lambda_1}+) = 0$. Using the ODEs (21) and (23), together with value matching and the definition of $\hat{p}_2$, establishes smooth pasting at $\hat{p}_2$. This implies $v_2'(\hat{p}_2) < 0$, and therefore $C_2^{rr} > 0$. Thus, $v_2$ is continuously differentiable, strictly decreasing and concave on $(p_1^*, \hat{p}_2)$ and convex on $(\hat{p}_2, 1)$.

We now show that $\hat{p}_2$ is well-defined, i.e. that there exists a unique $\hat{p}_2 \in (p_1^*, 1)$ such that

$$v_2^{rs}(\hat{p}_2) = \frac{\lambda_0(\lambda_1+\lambda_2) - \lambda_1\lambda_2\hat{p}_2}{r\lambda_2 + \lambda_0(\lambda_1+\lambda_2)}$$

At $p = p_1^*$, $v_2^{rs}(p) = \frac{\lambda_0}{r+2\lambda_0}$ and $\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} = \frac{\lambda_0\lambda_1}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Thus, we have

$$\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} - v_2^{rs}(p) = \frac{\lambda_0(r+\lambda_0)(\lambda_1-\lambda_2)}{(r\lambda_2+\lambda_0(\lambda_1+\lambda_2))(r+2\lambda_0)} > 0.$$

At $p = 1$, we have $v_2^{rs}(p) = \frac{\lambda_0}{r+\lambda_0+\lambda_1}$ and $\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} = \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Thus, we have

$$\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} - v_2^{rs}(p) = -\frac{\lambda_1(\lambda_2-\lambda_0)(r+\lambda_1)}{(r+\lambda_0+\lambda_1)(r\lambda_2+\lambda_0(\lambda_1+\lambda_2))} < 0.$$

For $p > p_1^*$, both $v_2^{rs}$ and $\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ are decreasing in $p$. The slope of $v_2^{rs}$ is bounded below by $-\frac{\lambda_0\lambda_1}{(r+\lambda_0)(r+\lambda_1+\lambda_0)}$, while the slope of $\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ is $-\frac{\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Since

$$\frac{\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} - \frac{\lambda_0\lambda_1}{(r+\lambda_0)(r+\lambda_1+\lambda_0)} = \frac{\lambda_1[r\lambda_2(r+\lambda_0+\lambda_1)+\lambda_0\lambda_1(\lambda_2-\lambda_0)]}{(r\lambda_2+\lambda_0(\lambda_1+\lambda_2))[(r+\lambda_0)(r+\lambda_0+\lambda_1)]} > 0,$$

there exists a unique $\hat{p}_2 \in (p_1^*, 1)$ such that

$$v_2^{rs}(\hat{p}_2) = \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2\hat{p}_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}.$$

By our discussion of best responses in the text, smooth pasting at $\hat{p}_2$ implies that $\lambda_2\hat{p}_2(1-v_2(\hat{p}_2)-(1-\hat{p}_2)v_2'(\hat{p}_2)) = \lambda_0(1-v_2(\hat{p}_2))$. As $v_2'(\hat{p}_2) < 0$, this implies $\hat{p}_2 < \frac{\lambda_0}{\lambda_2}$.

It remains to show that our payoff functions satisfy the Bellman equation (4). First, consider the range $[0, \frac{\lambda_0}{\lambda_1}]$. As $v_i = \frac{\lambda_0}{r+2\lambda_0}$ and $v_i' = 0$ in this range, it is immediate that $k_i = 0$ solves the Bellman equation in this range.

Next, let us consider the range $(\frac{\lambda_0}{\lambda_1}, \hat{p}_2]$. As $v_1 > \frac{\lambda_0}{r+2\lambda_0}$ in this range, $k_1 = 1$ satisfies the Bellman equation. Since $v_2(p) \leq \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ for all $p \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2]$, our discussion of best responses immediately implies that $k_2 = 0$ satisfies the Bellman equation as well.

Finally, we consider the range of beliefs $(\hat{p}_2, 1]$. By convexity of $v_2$ and smooth pasting at $\hat{p}_2$, $v_2' > \frac{-\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ in this range. This implies that $v_2(p) > \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$

for all $p$ in this range, and hence player 2 is playing a best response.

To show the best-response property for player 1, we consider the function $\tilde{v_1}(p) = \frac{\lambda_1}{r+\lambda_1+\lambda_2}p + \tilde{C}\mu(p)$, where the constant $\tilde{C}$ is implicitly defined by $\tilde{v_1}(\hat{p_2}) = \frac{\lambda_0}{r+2\lambda_0}$. From (23) it follows that $\tilde{v_1}'(p) > 0$ if and only if $p > \frac{r\lambda_0}{r\lambda_1+\lambda_0(\lambda_1-\lambda_2}$. As $\hat{p_2} > \frac{\lambda_0}{\lambda_1} \geq \frac{r\lambda_0}{r\lambda_1+\lambda_0(\lambda_1-\lambda_2)}$, we can conclude that $\tilde{v_1} > \frac{\lambda_0}{r+2\lambda_0}$ for all $p > \hat{p_2}$. Since $v_1^{rr}(\hat{p_2}) > \tilde{v_1}(\hat{p_2})$ and $v_1^{rr}(1) = \tilde{v_1}(1)$, we can conclude that $v_1^{rr}(p) > \tilde{v_1}(p)$, and hence that player 1 is playing a best response as well, for all $p \in (\hat{p_2}, 1)$.

## E.2 Uniqueness

Let $((k_1(p), k_2(p))_{p\in[0,1]}$ be an equilibrium of the game and define $p_l = \inf\{p \in [0,1] : \exists i \in \{1,2\}, k_i = 1\}$. If $p_l > \frac{\lambda_0}{\lambda_1}$, firm 1 has profitable deviation on $(\frac{\lambda_0}{\lambda_1}, p_l)$. Thus, $p_l \leq \frac{\lambda_0}{\lambda_1}$.

Suppose that $p_l < \frac{\lambda_0}{\lambda_1}$. There are now two possibilities. ($i$) First, suppose both firms are using $R$ to the immediate right of $p_l$. Note that, for any $p < \frac{\lambda_0}{\lambda_1}$, we have $\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} > \frac{\lambda_0\lambda_1}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2} > \frac{\lambda_0}{r+2\lambda_0} > 0$. As payoffs are continuous and both firms' payoff at $p = p_l$ is equal to $\frac{\lambda_0}{r+2\lambda_0}$, we will have

$$\frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} > v_2(p)$$

in some right-neighbourhood of $p_l$, and thus firm 2 is not playing a best response–a contradiction. Thus, suppose that ($ii$) only one of the firms, firm $i$, is using $R$ at beliefs just above $p_l$. As $p_l < \frac{\lambda_0}{\lambda_1} < \frac{\lambda_0}{\lambda_2}$, (19) implies that $v_i' < 0$ for beliefs just above $p_l$. This implies that $v_i$ drops below $\frac{\lambda_0}{r+2\lambda_0}$ in some right-neighbourhood of $p_l$, implying that firm $i$ is not playing a best response there. We thus conclude that $p_l = \frac{\lambda_0}{\lambda_1}$.

We will now establish that, in any equilibrium, there exists a right-neighbourhood of $\frac{\lambda_0}{\lambda_1}$ in which firm 1 plays $R$ while firm 2 plays $S$. First, suppose to the contrary that both firms play $R$ just above $\frac{\lambda_0}{\lambda_1}$. Then, by the same argument as above, $v_2 < \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ for some beliefs just above $\frac{\lambda_0}{\lambda_1}$, implying that firm 2 is not playing a best response there. By the same token, it is not possible that only firm 2 uses

$R$ in equilibrium to the immediate right of $\frac{\lambda_0}{\lambda_1}$, because, by (19), the payoff of firm 2 would fall below $\frac{\lambda_0}{r+2\lambda_0}$–a contradiction. We have thus established that, in any equilibrium, firm 1 will play $R$ while firm 2 will play $S$ in some right-neighbourhood of $\frac{\lambda_0}{\lambda_1}$.

Next, we will argue that, in no equilibrium, there exists a $p^{'} \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2)$ such that to the immediate right of $p^{'}$, firm 2 uses the method $R$ and firm 1 uses $S$. Suppose to the contrary that such a $p'$ exists and let $p_l^{'}$ be the lowest of such beliefs $p'$. Then, the payoff function of firm 2 (20) is strictly less than $\frac{\lambda_0}{r+2\lambda_0}$ to the immediate right of $p_l^{'}$, implying that firm 2 is not playing a best response.

By the same token, let $p_l''$ be the lowest belief in $(\frac{\lambda_0}{\lambda_1}, \hat{p}_2)$ such that both firms use method $S$ in some right-neighbourhood of $p_l''$. As $p_l'' > \frac{\lambda_0}{\lambda_1}$, firm 1's payoff $v_1(p_l''-) > \frac{\lambda_0}{r+2\lambda_0}$, implying firm 1 has a profitable deviation.

Now, let $p_l'''$ be the lowest belief in $(\frac{\lambda_0}{\lambda_1}, \hat{p}_2)$ such that both firms use method $R$ in some right-neighbourhood of $p_l'''$. Then, firm 2's payoff satisfies $v_2(p_l''') = v_2^{rs}(p_l''') < \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$, where the inequality follows from $p_l''' < \hat{p}_2$, implying that firm 2 has a profitable deviation.

We have thus established that, in any equilibrium, for $p \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2)$, firm 1 uses $R$ and 2 uses $S$. We will now argue that for all $p > \hat{p}_2$, using method $R$ is the dominant action for firm 1. Suppose not and let $\tilde{p}$ be the lowest belief in $(\hat{p}_2, 1)$ such that firm 1 uses $S$ while firm 2 uses $R$ in some right-neighbourhood of $\tilde{p}$. From our verification arguments in Appendix (E.1), we can argue that firm 1 is not playing a best response at beliefs just above $\tilde{p}$. A similar argument to above furthermore establishes that either firm would have a profitable deviation at the lowest belief $\tilde{p}' \in (\hat{p}_2, 1)$ such that both firms use $S$ is some right-neighbourhood of $\tilde{p}'$. This shows that for all $p > \hat{p}_2$, using method $R$ is the dominant action of firm 1. From the equilibrium constructed in the preceding proposition, it follows that the unique best response of firm 2 is to choose $R$, which concludes the proof.

# F Proof of Proposition 6

We prove this proposition by the following two lemmata.

**Lemma 1** *There exists a $\tilde{r} \in (r_0, \infty)$ such that for all $r > \tilde{r}$, we have $p_2^*(r) < p_2^*(r_0)$.*

**Proof.** Define $\tilde{p}(r) = \frac{\lambda_0(r+\lambda_0+\lambda_1)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. We note that $\tilde{p}'(r) < 0$ and $\lim_{r\to\infty} \tilde{p}(r) = \frac{\lambda_0}{\lambda_2}$. The defining equation of $p_2^*$, together with the fact that $C_{rs} > 0$ in the planner's value function (3), implies that $p_2^*(r) < \tilde{p}(r)$ for all $r > 0$. For all $0 < r < \infty$, we have $p_2^*(r) > \frac{\lambda_0}{\lambda_2}$ by Proposition 3. Thus, there exists a $\tilde{r}^p \in (0, \infty)$ such that, for all $r > \tilde{r}^p$, we have $p_2^*(r) < \tilde{p}(r) < p_2^*(r_0)$. ■

**Lemma 2** *There exists a $\check{r} \in (r_0, \infty)$ such that for all $r > \check{r}$, we have $\hat{p}_2(r) > \hat{p}_2(r_0)$.*

**Proof.** Define $\check{p}(r) = \frac{\lambda_0}{\lambda_2}\frac{r\lambda_1+\lambda_0(\lambda_1+\lambda_2)}{\lambda_1(r+2\lambda_0)}$. We note that $\check{p}'(r) > 0$ and $\lim_{r\to\infty} \check{p}(r) = \frac{\lambda_0}{\lambda_2}$. The defining equation of $\hat{p}_2$, together with the fact that $p \mapsto \frac{\lambda_0(\lambda_1+\lambda_2)-\lambda_1\lambda_2 p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} - v_2^{rs}(p)$ admits of a unique root in $(p_1^*, 1)$, implies that $\check{p}(r) < \hat{p}_2(r)$ for all $r > 0$. For all $0 < r < \infty$, we have $\hat{p}_2(r) < \frac{\lambda_0}{\lambda_2}$ by Proposition 4. Thus, there exists a $\check{r} \in (0, \infty)$ such that, for all $r > \tilde{r}^p$, we have $\hat{p}_2(r) > \tilde{p}(r) > \hat{p}_2(r_0)$. ■

Let $\bar{r} = \max\{\tilde{r}, \check{r}\}$. Then, for all $r > \tilde{r}$, $p_2^*(r) < p_2^*(r_0)$ and $\hat{p}_2(r) > \hat{p}_2(r_0)$.