DO STRONGER PATENTS LEAD TO FASTER INNOVATION? THE EFFECT OF DUPLICATIVE SEARCH*

Kaustav Das† Nicolas Klein‡

March 9, 2020

Abstract

We analyse a model of two firms that are engaged in a patent race. 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 coincides with the cartel solution if and only if firms are symmetric in their abilities of leveraging a good innovative method or there is no patent protection. Otherwise, equilibrium will entail excessive duplication of efforts in the innovative method, as compared to the cartel benchmark, for any level of patent protection. We show that the expected time to a breakthrough is minimised at an interior level of patent protection, providing a possible explanation for the decrease in R&D productivity sometimes associated with stronger patent protections.
JEL Classification Numbers: C73, D83, O31.

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

1 Introduction

Innovative methods are 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

*This paper supersedes two earlier versions entitled "Inefficient Duplicative Search in a Patent Race" and "Excessive Duplicative Search in a Patent Race."

†University of Leicester, School of Business, Brookfield, Leicester LE2 1RQ; Email: daskaustav84@gmail.com

‡Corresponding Author: Université de Montréal and CIREQ, Dpartement de Sciences conomiques, C.P. 6128 succursale Centre-ville Montral, QC H3C 3J7, Canada, Email:kleinnic@yahoo.com, Tel: (+1)514-343-7908

is characterised by both a decrease in acetylcholine (neurotransmitter) levels in the brain and the accumulation of β-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 β-amyloid plaques. However, 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 β-amyloid plaques via an increase in neurotransmitter levels.[1] Indeed, there is some evidence that Donepezil has a beneficial effect on the level of β-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. The importance of the latter depends on the level of patent protection afforded by the legal system. Concurrently, 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 optimal future 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[3] in a setting in which two firms are engaged in a patent race and their research choices are observable, using a variant of the two-armed *exponential-bandits* framework of Keller, Rady and Cripps (2005). We take the *scale* of R&D as given, analysing the problem of allocating a *given* resource flow among two competing methods of R&D. More specifically, 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 β-amyloid plaques, this would correspond to the search for a drug that seeks to fight the concentration of β-amyloid plaques by increasing 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 ends the game, yielding a payoff that is shared between the firms as a function of the level of patent protection that prevails. 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

[1] See Moss (2018).

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

[3] While the phrase *process innovation* is often used to refer to innovations that decrease the cost of production of a certain good, we here use the term to signify the use of a new method to achieve a certain goal.

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. There is, however, also a (negative) payoff externality between the firms, the strength of which depends on the level of patent protection afforded to the winner of the R&D race.

We first analyse the problem of a cartel that endeavours to maximise the aggregate expected discounted payoff. In the optimal cartel solution, the stronger firm experiments with the innovative method if and only if the firms' belief that the innovative method is good exceeds its myopic cutoff; i.e., if and only if the *instantaneous* expected arrival rate of a success is higher with the innovative method than the established one for the stronger firm. If the other firm is equally productive with the innovative method, it behaves exactly as the first firm in the cartel solution. Otherwise, the less productive firm anticipates that the more productive firm will continue exploring the innovative method until its 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 the cartel will apply a cutoff more optimistic than its myopic cutoff to the less productive firm. As the expected time to a breakthrough is minimised when both firms switch methods at their respective myopic cutoffs, this implies that a profit-maximising cartel will delay the expected breakthrough time by making the weaker firm give up on the innovative method prematurely.

We go on to show that, for any level of patent protection, 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*. If and only if both firms are equally productive with a good innovative work method or there is no patent protection whatsoever, the unique Markov perfect equilibrium coincides with the cartel solution. In the case of homogeneous firms, it also minimises the expected time to a breakthrough.

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 an excessive duplication of innovative efforts, as compared to the cartel solution. The stronger firm always acts as in the cartel benchmark, being the last to give up experimenting with the innovative method in equilibrium. The less productive firm, by contrast, trades off two effects: on the one hand, because of discounting, it wishes for a breakthrough to occur as quickly as possible, while, on the other hand, it would rather it was the one achieving the breakthrough than its competitor. Which of these effects prevails at its myopic belief threshold depends on the strength of the patent protection afforded to the first innovator. If patent protection, and thus the firms' payoff rivalry, is strong, the above reasoning explains why the less productive firm will endeavour to "eat up" some of the stronger firm's comparative advantage, by extending experimentation below its myopic cutoff, thereby reducing the expected time the stronger firm will spend on the innovative method. This implies that, for strong patent protections, research productivity, as measured by the expected time to a breakthrough, is *lowered* because the weaker firm extends research on the innovative method beyond its myopic threshold.

For weaker levels of patent protection, by contrast, the weaker firm will stop experimentation at a threshold exceeding its myopic cutoff, while still extending it beyond the cartel threshold. Thus, while the effect is mitigated as compared to the cartel, research productivity is still suboptimal because the weaker firm gives up on the innovative method too quickly. There is a unique, interior, level of patent protection that makes the weaker firm behave in equilibrium exactly as though it were myopic; this level of patent protection minimises the expected time to a breakthrough.

In summary, the cartel asks the weaker firm to step aside somewhat for the stronger firm in the cooperative solution. In equilibrium, by contrast, the weaker firm always stands aside less. If patent protection is strong, it is willing to incur a reduction in its own breakthrough rate by extending experimentation below its myopic threshold, in order to discourage the stronger firm by its hapless experimentation. If patent protection is weak, the weaker firm is incurring a reduction in its own breakthrough rate by switching to the established method at a threshold above its myopic cutoff, in order to encourage the stronger firm to spend more time exploring the innovative method.

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 1990-2010, 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. Incidentally, this time period coincides with the implementation of the Agreement on Trade Related Intellectual Property Rights (TRIPS), which covers pharmaceutical products or processes invented since January 1, 1995. This agreement obliges all WTO members to afford patent protections for pharmaceutical inventions. Previously to the TRIPS agreement, copies of medicines that were patent-protected elsewhere were often widely available in many developing countries.[4] TRIPS has thus significantly strengthened patent protections in the pharmaceutical sector.

Our model thus suggests a novel effect possibly explaining the decrease in pharmaceutical R&D productivity connected with stronger patent protections. Indeed, studying pharmaceutical patent pro-

[4] See e.g.Boulet et al. (2000).

tection for the time period 1978–2002, Qian (2007) writes that "there appears to be an optimal level of intellectual property rights regulation above which further enhancement reduces innovative activities" (see his Abstract). He goes on to note: "National patent laws would also induce domestic investors to switch from imitative activities to innovative ones" (see Qian (2007), p. 437). Sakakibara and Branstetter (2001) likewise find no evidence of an increase in R&D output subsequent to a strengthening of patent projections in Japan in 1988.[5] Our model formalises the idea that excessive duplication of innovative efforts may be generated in an R&D race with observable actions as a result of strong patent protections.

Related Literature:

The only paper we are aware of that provides a theoretical explanation for a negative link between patent strength and reduced R&D investment and output is Bessen and Maskin (2009), which assumes that innovations are both sequential and complementary. 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. Schneider and Wolf (2019) analyses the problem of a single agent who faces a deadline to solve a given problem and dynamically chooses between an innovative, and potentially quick, method and a traditional, slower, method.

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 akin to that 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. 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.

Besanko and Wu (2013) also introduces payoff externalities into a Keller, Rady and Cripps (2005) setting. Their paper differs from ours, inter alia, in that their players are ex ante homogeneous and 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.[6] Akcigit and Liu (2016) also anal-

[5]They find no evidence of an increase in R&D expenditure either. This is consistent with our model if we interpret investment in the traditional method as R&D expenditure as well. Bessen and Hunt (2004) also find that R&D intensity in the US software industry decreased subsequently to an enhancement of patent protections for computer programs that occurred in the US during the 1980s and 1990s by virtue of a gradual evolution of the appertaining jurisprudence. Our model only pertains to the choice of methodology in R&D races, rather than the level of R&D investments.

[6]In a different context, Negoescu et.al (2018) adapt the framework of Keller, Rady and Cripps (2005) in a medical

yses a variant of Keller, Rady and Cripps (2005)'s two-armed bandit model with one risky and one safe method of investigation, while assuming players cannot return to an arm they have previously discarded. In Akcigit and Liu (2016), the risky arm yields either a publicly observable *breakthrough*, if good, or a *breakdown*, which is only privately observed, if bad, at the first jumping time of a Poisson process whose rate does not depend on the state of the risky project. A firm may refrain from disclosing such a breakdown, as it benefits from its competitor's resources being devoted to a dead end. In Akcigit and Liu (2016)'s model, inefficiencies can arise from both wasteful dead-end replication and from an early abandonment of the risky project on account of the positive informational externality.

In a model without informational externalities, Wong (2018) analyses an R&D race where firms choose when irreversibly to exit and innovation can occur through a technology of initially unknown quality only; this quality is independent across firms. If the prior probabilities of the risky technology being good differ between the firms, equilibrium under winner-takes-all payoffs involves too much experimentation compared to the solution which maximises the aggregate payoffs. There exist (weaker) levels of patent protection that lead to higher joint payoffs in equilibrium. However, when prior probabilities differ, there is no level of patent protection that can implement the cartel solution; indeed, for any given level of patent protection, either the pessimistic firm quits too late or the optimistic firm quits too soon, compared to the cartel solution. In our model, by contrast, the quality of the risky project is always identical between firms, so that the positive informational externality familiar from strategic-experimentation models à la Keller, Rady and Cripps (2005) arises; heterogeneity in our model pertains to the firms' abilities of achieving a breakthrough *conditionally* on the risky method being good. Moreover, firms choose a *method* of pursuing innovation, rather than *whether* to innovate at all; innovation therefore eventually takes place almost surely. Finally, in our setting, the firms' aggregate payoffs are maximised in equilibrium when there is no patent protection at all, while Wong (2018) finds under-experimentation in the absence of patent protections. In our setting, an interior level of patent protection minimises the equilibrium time to breakthrough.

Halac, Kartik and Liu (2016) characterises sharing and disclosure rules in contests that maximise the probability of an innovation, the feasibility of which is ex-ante uncertain. Bimpikis, Ehsani and Mostagir (2019) considers two-stage contests and analyses the mechanisms for disclosure of a first-stage breakthrough a designer will want to commit to. We, by contrast, abstract from both private information and the question of effort provision to focus on firms' choice of method; moreover, a breakthrough consists of a single successful stage.

Our paper also contributes to the relatively less explored area of choice of methodological approach in R&D races.[7] Indeed, we show that, when firms are heterogeneous and there is *some*,

decision-making setting to assess drug effectiveness through clinical trials.

[7]As mentioned above, we here abstract from the choice of the *scale* of R&D, to focus our analysis on the 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)).

however modest, level of patent protection, there is always some excessive duplication of innovative research effort compared to the cartel optimum. Whether there is excessive or insufficient duplication compared to the optimum of a social planner who wants to speed up as much as possible the arrival of a breakthrough depends on the strength of the patent regime. In a static model with winner-takes-all competition, Klette and de Meza (1986) allows firms to choose the riskiness of R&D strategies. They find that, if invention times are symmetrically distributed, the market equilibrium entails riskier R&D strategies than the social optimum. Also analysing a static model with winner-takes-all competition, Bhattacharya and Mookerjee (1986) shows 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 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. Choi and Gerlach (2014) considers the choice between an easier and a harder project, which are complementary. In contrast to our setting, the success probabilities for each project are common knowledge. They show that, in equilibrium, there tends to be excessive duplication of efforts in the easier project. Brian and Lemus (2017) also analyses the choice of research project, when the success probabilities for each project are common knowledge. The complementarity structure across projects is given by a general innovation graph. In equilibrium, firms tend excessively to duplicate efforts either in easy but relatively non-lucrative projects or in difficult projects that yield a big immediate payoff but do not entail a large continuation value. In order to guarantee that firms behave in a socially optimal way in Markov perfect equilibrium in their model, Brian and Lemus (2017) shows that it is necessary that the transfer policy condition on the properties of projects that are not discovered along the path of play. 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. Section 5 concludes. Most proofs are in the Appendix.

2 The Environment

Two firms are simultaneously trying to achieve a breakthrough in continuous time. The first breakthrough yields a payoff of α to the firm accomplishing it, and $(1-\alpha)$ to the competing firm where $\alpha \in [\frac{1}{2}, 1]$. We interpret the parameter α as measuring the strength of patent protection afforded to the firm achieving the first breakthrough. Indeed, $\alpha > \frac{1}{2}$ implies that the firm accomplishing the first breakthrough gets a premium, with $\alpha = 1$ corresponding to the *winner-takes-all* case; 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, method S, is *established* (*safe*) in that it yields a breakthrough at the first jumping time of a Poisson process with known intensity $\lambda_0 > 0$. The other method, method R, 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, 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 cartel's problem, which seeks to maximise the firms' aggregate discounted payoffs.

3.1 The cartel's problem

Without loss of generality, we can restrict the cartel to Markov strategies $k(p_t)$ with the posterior belief p_t as the state variable, where k denotes the number of firms the cartel assigns to method R.[8] This implies $k(p_t) \in \{0,1,2\}$. Let $v(p)$ be the value function of the cartel. Then we have

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

[8]We suppress the arguments whenever this is convenient.

The expression $2\lambda_0(1-v)$ denotes the expected flow payoff the cartel can guarantee itself by using method S. The expression $\lambda_1 p\left(1-v-(1-p)v'\right)-\lambda_0(1-v)$ reflects the premium the cartel 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 cartel to choose $\{k(p_t)\}_{t\geq 0}$ with $k(p_t)\in\{0,2\}$ for all $t\geq 0$. The cartel's solution is described in the following proposition. It shows that maximisation of joint profits 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, $\mu(p)=(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{2\lambda_1}}$.

Proposition 1 *The cartel'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 cartel's value function is given by

$$v(p)=\begin{cases} \frac{2\lambda_1 p}{r+2\lambda_1}+\frac{4\lambda_0(\lambda_1-\lambda_0)\mu(p)}{(r+2\lambda_0)(r+2\lambda_1)\mu(p_1^*)} & \text{if } p\in(p_1^*,1] \\ \frac{2\lambda_0}{r+2\lambda_0} & \text{if } p\in[0,p_1^*] \end{cases}$$

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 cartel's value. ■

Note that, in the homogeneous case, the cartel's policy minimises the expected time to a breakthrough. 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

$$\begin{aligned} rv_i = \max_{k_i\in\{0,1\}} & \lambda_0[1-2v_i]+k_i\{\lambda_1 p[\alpha-v_i-(1-p)v_i']-\lambda_0[\alpha-v_i]\} \\ & +k_j\{\lambda_1 p[(1-\alpha)-v_i-(1-p)v_i']-\lambda_0[(1-\alpha)-v_i]\}. \end{aligned} \tag{1}$$

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

$$v_i \geq \frac{2\lambda_0\alpha + \lambda_1 p(1-2\alpha)}{r+2\lambda_0}$$

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 cartel's solution.

Proposition 2 *If firms are homogeneous, the unique MPE coincides with the cartel's solution (and thus minimises the expected time to a breakthrough), for any level of patent protection* $\alpha \in [\frac{1}{2}, 1]$.

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 cartel, which seeks to maximise the firms' aggregate discounted payoffs.

4.1 The Cartel's Problem

We can again restrict the cartel 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 cartel assigns firm i to method R (S). The value function of the cartel $v(p)$ satisfies

$$\begin{aligned} rv = \max_{(k_1,k_2)\in\{0,1\}^2} & 2\lambda_0(1-v) + k_1\{\lambda_1 p[1-v-(1-p)v^{'}] - \lambda_0[1-v]\} \\ & + k_2\{\lambda_2 p[1-v-(1-p)v^{'}] - \lambda_0[1-v]\} \end{aligned} \tag{2}$$

The expression $2\lambda_0(1-v)$ is the expected flow payoff the cartel can guarantee itself by using the method S. On the other hand, $\lambda_i p[1-v-v^{'}(1-p)] - \lambda_0(1-v)$ reflects the premium the cartel gets by assigning firm i to method R. By linearity, it would be without loss for the cartel 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 cartel's solution. It shows that maximisation of joint payoffs requires the cartel to choose the myopically optimal method for firm 1, while assigning firm 2 to method S at some beliefs above 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 cartel'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 cartel's value function is given by*

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

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

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

C_{rs} and C_{rr} are constants of integration with $C_{rs} = \frac{r\lambda_0(\lambda_1-\lambda_0)}{(r+\lambda_0)(r+2\lambda_0)(r+\lambda_0+\lambda_1)\mu_1(p_1^*)} > 0$, and $C_{rr} > 0$ is determined from $\check{v}_{rr}(p_2^*) = \check{v}_{rs}(p_2^*)$.

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 cartel's value function. ■

The cartel'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 cartel 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 cartel can be intuitively explained as follows. Since the game ends after the first breakthrough, there is no learning benefit from a breakthrough and hence, in the cartel'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 cartel would gain if it 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 cartel to substitute *future* experimentation by firm 1 for *current* experimentation by firm 2. This intertemporal substitution, of course, comes at the price of delaying the expected time of the breakthrough. 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 cartel incurs an expected loss by giving up current experimentation by firm 2. At $\frac{\lambda_0}{\lambda_2}$, the myopic threshold belief of firm 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 firm 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 firm 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 from firm 2 using R, $(\lambda_2 p - \lambda_0)(1-v) = 0$, is outweighed by the cost $-\lambda_2 p(1-p)v^{'} < 0$. The cartel's incentive to substitute future experimentation by the stronger firm for current experimentation by the weaker firm thus leads to a delay in the expected time of breakthrough, suggesting that collusion between heterogeneous firms harms their research productivity.

4.2 Non-cooperative game

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

$$v_i = \max_{k_i \in \{0,1\}} \{\lambda_0(1-k_i)\alpha\, dt + \lambda_0(1-k_j)(1-\alpha)\, dt + k_i\lambda_i p\alpha\, dt + k_j\lambda_j p(1-\alpha)\, dt$$

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

$$\Rightarrow rv_i = \lambda_0[1-2v_i] + \max_{k_i \in \{0,1\}} k_i\{\lambda_i p[\alpha - v_i - (1-p)v_i^{'}] - \lambda_0[\alpha - v_i]\}$$
$$+ k_j\{\lambda_j p[1-\alpha - v_i - (1-p)v_i^{'}] - \lambda_0[1-\alpha - v_i]\}. \quad (4)$$

Firm i ($i = 1,2$) can guarantee itself an expected flow payoff of $\lambda_0(\alpha - v_i) + ((1-k_j)\lambda_0 + k_j\lambda_j p)(1-\alpha - v_i) = \lambda_0(1-2v_i) + k_j(\lambda_j p - \lambda_0)(1-\alpha - v_i)$ by using the traditional method (S). The term $\{\lambda_i p[\alpha - v_i - (1-p)v_i^{'}] - \lambda_0[\alpha - v_i]\}$ captures the premium firm i receives by using the innovative method. The expression $((1-k_j)\lambda_0 + k_j\lambda_j p)(1-\alpha - v_i)$ captures the payoff externality firm j exerts on firm i if it has a success, while $k_j\lambda_j p(1-p)v_i'$ captures the informational externality caused by firm j's hapless experimentation with the risky method.

Best Responses:

Suppose $k_j = 0$ ($j \in \{1,2\}$) in an open neighbourhood of p. From (4) and ODE (21), we can see that using method R in a neighbourhood 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 neighbourhood.

Next, suppose $k_j = 1$ in an open neighbourhood of p. From (4) and ODE (25), we can infer that choosing R is optimal for firm i in a neighbourhood of p if and only if

$$v_i \geq \frac{\lambda_0 \alpha[\lambda_1+\lambda_2]+\lambda_1\lambda_2 p[1-2\alpha]}{r\lambda_i+\lambda_0(\lambda_1+\lambda_2)} \tag{5}$$

is satisfied in that neighbourhood.

Our main result characterizes the unique Markov perfect equilibrium of our game. For any level of patent protection α, both firms will use a cutoff strategy in equilibrium, that is, they use the innovative method if and only if the likelihood of it being good is above a threshold. 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 (R) on $(\hat{p}_2(\alpha), 1]$ and the safe method (S) otherwise. Thus, while firm 1's cutoff is independent of α, firm 2's equilibrium threshold is a decreasing function of α; the stronger the level of patent protection α, the more firm 2 will be inclined to use the innovative method. Indeed, the theorem shows that *firm 2 will use the risky method too much*, compared to the cartel solution, as soon as there is *some* level of patent protection, i.e., whenever $\alpha > \frac{1}{2}$. If patent protection is relatively weak, i.e., $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$, firm 2 will use the risky method less than if it were by itself; for $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, by contrast, it uses the risky method beyond the myopically optimal threshold. If and only if $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$ will it behave myopically, thereby minimising the expected time to a first breakthrough.

This can be intuitively understood as follows. Firms have two goals: (1) on account of discounting, they want the breakthrough to occur as soon as possible; (2) on account of the payoff rivalry between them, they both want to be the one achieving the breakthrough. The level of patent protection determines the relative importance of these goals in the firms' objectives. When there is no patent protection, i.e., $\alpha = \frac{1}{2}$, the payoff rivalry is shut down and firms behave cooperatively in the unique Markov perfect equilibrium of the non-cooperative game; i.e., $\hat{p}_2(\frac{1}{2}) = p_2^*$. As soon as $\alpha > \frac{1}{2}$, some payoff rivalry comes into play, as both firms want to be the first inventor achieving the breakthrough; as a result, $\hat{p}_2(\alpha) < p_2^*$ for all $\alpha > \frac{1}{2}$. At the belief $p = \frac{\lambda_0}{\lambda_2}$, the individual myopic expected 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 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 $p = \frac{\lambda_0}{\lambda_2}$, this reduces both the overall chance of a breakthrough on the risky method and, more particularly, the chances of a breakthrough by firm 1. While the former is bad news for firm 2, the latter is good news (provided $\alpha > \frac{1}{2}$); now, how firm 2 trades-off these two countervailing effects depends on the level of patent protection α. For $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$, the two effects just cancel out and firm 2 best-responds by behaving myopically, i.e., as though it were by itself. For high levels of patent protection ($\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$), the desire to be first dominates at the myopic threshold and firm 2 extends the use of the risky method below its myopic threshold, while for low levels of patent protection ($\alpha \in (\frac{1}{2}, \frac{r+\lambda_0}{r+2\lambda_0})$), firm 2's cooperative motive prevails at the myopic threshold, and its equilibrium cutoff satisfies $\hat{p}_2(\alpha) \in (\frac{\lambda_0}{\lambda_2}, p_2^*)$.

Theorem 1 *There exists a unique Markov perfect equilibrium. Equilibrium strategies are given by* $k_1^{-1}(1) = (p_1^*, 1]$ *and* $k_2^{-1}(1) = (\hat{p}_2(\alpha), 1]$. *Firm 2's cutoff* $\hat{p}_2$ *is a strictly decreasing, continuously differentiable, function, satisfying* $\hat{p}_2(\frac{1}{2}) = p_2^*$, $\hat{p}_2(\frac{r+\lambda_0}{r+2\lambda_0}) = \frac{\lambda_0}{\lambda_2}$, *and* $\hat{p}_2(1) > p_1^*$.

The firms' equilibrium payoffs are given by

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

and

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

respectively.

The threshold $\hat{p}_2(\alpha)$ *is implicitly defined by* $v_2^{rs}(\hat{p}_2(\alpha)) = \frac{\lambda_0\alpha[\lambda_1+\lambda_2]+\lambda_1\lambda_2 p[1-2\alpha]}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. *The constants of integration are determined by value matching, i.e.,* $C_1^{rs} > 0$ *is given by* $v_1^{rs}(p_1^*) = \frac{\lambda_0}{r+2\lambda_0}$ *and* C_2^{rs} *by* $v_2^{rs}(p_1^*) = \frac{\lambda_0}{r+2\lambda_0}$. *We have* $C_2^{rs} \geq 0$ ($C_2^{rs} \leq 0$) *if and only if* $\alpha \leq \frac{r+\lambda_0}{r+2\lambda_0}$ ($\alpha \geq \frac{r+\lambda_0}{r+2\lambda_0}$), *with* $C_2^{rs} = 0$ *if and only if* $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$. *Similarly, the constants of integration* C_1^{rr} *and* $C_2^{rr} > 0$ *are determined by* $v_1^{rr}(\hat{p}_2(\alpha)) = v_1^{rs}(\hat{p}_2(\alpha))$, *and* $v_2^{rr}(\hat{p}_2(\alpha)) = v_2^{rs}(\hat{p}_2(\alpha))$, *respectively. The function* v_2 *is smooth, while* v_1 *is smooth everywhere except at* $p = \hat{p}_2(\alpha)$.

Proof. Existence follows from standard verification arguments, while uniqueness follows from the Bellman equation (4) and the relevant ODEs (Appendix A.2). Please see Appendix E for a detailed proof. ■

For high values of patent protection ($\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$), firm 2's value function is decreasing and concave in the region where only firm 1 uses method R; it is convex in the range where both firms use it. It has an inflection point at $\hat{p}_2(\alpha)$, where firm 2 switches methods, and eventually becomes increasing as firms become very optimistic about method R. For low levels of patent protection ($\alpha \leq \frac{r+\lambda_0}{r+2\lambda_0}$), by contrast, v_2 is increasing and convex throughout.

It seems reasonable to assume that a social planner may want to speed up as much as possible the arrival time of a breakthrough, such as a cure for Alzheimer's disease. Our analysis would suggest that, in the knife-edge case of perfectly homogeneous firms, both the cartel and the non-cooperative firms would behave consistently with this goal. However, as soon as one firm is better capable of handling the innovative method, the cartel steers the less productive firm away from the innovative method too soon, i.e., there is *insufficient duplication* in the search for an innovative method. What happens in non-cooperative equilibrium in the heterogeneous case depends on the level of patent protection, as summarised in the following corollary.

Corollary 1 *Suppose the firms are heterogeneous, i.e., $\lambda_1 > \lambda_2$. The expected time to the first breakthrough is minimised for the level of patent protection $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$. If patent protection is strong, i.e., $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, this expected time is delayed on account of* excessive duplication *of innovative efforts. If patent protection is weak, i.e., $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$, this expected time is delayed on account of* insufficient duplication *of innovative efforts. The delay due to insufficient duplication is worst when there is no patent protection at all (i.e. $\alpha = \frac{1}{2}$) or firms form a cartel.*

Thus, our analysis would suggest that, besides watching out for collusion between firms, policymakers who endeavour to speed up the expected time of a decisive breakthrough should be wary of both too strong patent regimes as well as too weak ones. Indeed, the former will tend to exacerbate firms' rivalry to the point where the race to be first makes them engage in excessive duplication of innovative efforts. The latter, by contrast, makes firms behave "too cooperatively" in the sense that the weaker firm will be too inclined to substitute *future* experimentation by its partner for its own *current* experimentation, leading to insufficient duplication of research efforts. Our findings thus suggest a formal channel explaining the puzzling decrease in R&D productivity connected with a strengthening of patent protections, which has been noted in the empirical literature.

5 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 higher amounts of experimentation in equilibrium than in the cartel's solution. This effect is the stronger the more potent the regime of patent protection. The equilibrium expected time to breakthrough is minimised for an interior level of patent protection. This expected time to breakthrough is delayed on account of excessive (insufficient) duplication of experimentation if patent protection is above (below) this threshold.

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). In our setting, moreover, equilibrium deviates from the cooperative solution because of higher information production, while all equilibria in Keller, Rady and Cripps (2005) have the feature that players experiment too little compared to the cooperative benchmark.

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 of whether 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.

Acknowledgements

We thank Siddhartha Bandyopadhyay, Subir Bose, Dirk Bergemann, Kalyan Chatterjee, Martin Cripps, Sanjeev Goyal, Johannes Hörner, Vijay Krishna, Igor Letina, Arina Nikandrova, 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.

References

Akcigit U, Liu Q (2016). The Role of Information in Competitive Experimentation, *Journal of the European Economic Association*, 14(4), 828-870.

Besanko D, 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.

Bessen J, Hunt R (2004). An Empirical Look At Software Patents, *Working paper, Federal Reserve Bank of Philadelphia*.

Bessen J, Maskin E (2009). Sequential innovation, patents, and imitation, *The RAND Journal of Economics*, 40(4), 611-635.

Bhattacharya S, Mookerjee D (1986). Portfolio Choice in Research and Development, *Rand Journal of Economics*, 17, 594-605.

Bimpikis K, Ehsani S, Mostagir M (2019). Designing Dynamic Contests, *Operations Research*, 67(2), 339-356.

Bolton P, Harris C (1999). Strategic Experimentation, *Econometrica*, 67, 349-374.

Boulet P, Perriens J, Renaud-Thiery F, Velasquez G (2000). Pharmaceuticals and the WTO TRIPS Agreement: Questions and Answers, *UNAIDS/WHO document*.

Bryan K, Lemus J (2017). The direction of innovation, *Journal of Economic Theory*, 172, 247-272.

Chatterjee K, Evans R (2004). Rivals' Search for Buried Treasure: Competition and Duplication in R&D, *Rand Journal of Economics*, 35, 160-183.

Choi JP, Gerlach H (2014). Selection Biases in Complementary R&D Projects, *Journal of Economics and Management Strategy*, 23(4), 899-924.

Das K, Klein N, 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, 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, Csernansky J G (2009). Effects of Donepezil on Amyloid-β 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*.

Halac M, Kartik N, Liu Q (2016). Optimal Contracts for Experimentation, *The Review of Economic Studies*, 83(3), 1040-1091.

Keller G, Rady S, Cripps M (2005). Strategic Experimentation with Exponential Bandits, *Econometrica*, 73, 39-68.

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

Klette T, de Meza D (1986). Is the Market Biased Against Risky R&D?, *The RAND Journal of Economics*, 17(1), 133-139.

Lee T, 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.

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

Negoescu D, Bimpikis K, Brandeau M, Iancu D (2018). Dynamic Learning of Patient Response Types: An Application to Treating Chronic Diseases, *Management Science*, 64(8), 3469-3488.

Pammolli F, Magazzini L, Riccaboni M (2011). The Productivity Crisis in Pharmaceutical R&D, *Nature Reviews Drug Discovery*, 10(6), 428-438.

Qian Y (2007). Do National Patent Laws Stimulate Domestic Innovation In a Global Patenting Environment? A Cross-Country Analysis of Pharmaceutical Patent Protection, 1978-2002, *The Review of Economics and Statistics*, 89(3), 436-453.

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

Sakakibara M, Branstetter L (2001). Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms, *The RAND Journal of Economics*, 32(1), 77-100.

Schneider J, Wolf C (2019). Switching Forth and Back: Problem Solving under Time Pressure, *available at SSRN*.

Wong T (2018). Free Riding and Duplication in R&D, *Working paper, Aalto University*.

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

Cartel's problem:

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

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

This is solved by

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

The non-cooperative game:

If both firms adopt S, either firm's value function satisfies

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

If both firms adopt R on an open set of beliefs, either firm's value function satisfies

$$2\lambda_1 p(1-p)v' + (r + 2\lambda_1 p)v = \lambda_1 p \tag{10}$$

on this open set. This ODE is solved by

$$v(p) = \frac{\lambda_1 p}{r + 2\lambda_1} + C\mu(p). \tag{11}$$

Now, suppose firm i adopts method R and firm j adopts method S on an open set of beliefs. Then, firm i's value function satisfies

$$\lambda_1 p(1-p)v_i' + (r + \lambda_0 + \lambda_1 p)v_i = \lambda_0(1-\alpha) + \lambda_1 p\alpha \tag{12}$$

on this open set. This is solved by

$$v_i(p) = \frac{\lambda_0(1-\alpha)}{r + \lambda_0} + \frac{\lambda_1 p}{r + \lambda_0 + \lambda_1}\left[\alpha - \frac{\lambda_0(1-\alpha)}{r + \lambda_0}\right] + C\mu_i(p). \tag{13}$$

By the same token, j's value function satisfies

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

This is solved by

$$v_j(p) = \frac{\lambda_0 \alpha}{r+\lambda_0} + \frac{\lambda_1 p}{r+\lambda_0+\lambda_1}\left[1-\alpha-\frac{\lambda_0\alpha}{r+\lambda_0}\right] + C\mu_i(p). \tag{15}$$

A.2 ODEs in the game with asymmetric firms

Cartel's problem:

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

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

$$\lambda_1 p(1-p)v^{'} + (r+\lambda_0+\lambda_1 p)v = \lambda_0 + \lambda_1 p. \tag{16}$$

This is solved by

$$v(p) = \frac{\lambda_0}{r+\lambda_0} + \frac{r\lambda_1}{(r+\lambda_0)(r+\lambda_0+\lambda_1)}p + C\mu_1(p). \tag{17}$$

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

$$(\lambda_1+\lambda_2)p(1-p)v^{'} + (r+(\lambda_1+\lambda_2)p)v = (\lambda_1+\lambda_2)p. \tag{18}$$

This is solved by

$$v(p) = \frac{\lambda_1+\lambda_2}{r+\lambda_1+\lambda_2}p + C\mu(p). \tag{19}$$

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

$$\frac{\lambda_0}{r+2\lambda_0}. \tag{20}$$

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

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

The solution to the above differential equation is

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

Firm j's payoff satisfies

$$\lambda_i p(1-p)v_j' + (r+\lambda_0+\lambda_i p)v_j = \lambda_0\alpha + \lambda_i p(1-\alpha). \tag{23}$$

The solution to the above differential equation is

$$v_j^{rs}(p) = \frac{\lambda_0\alpha}{r+\lambda_0} + \frac{\lambda_i p}{r+\lambda_0+\lambda_i}\left[1-\alpha-\frac{\lambda_0\alpha}{r+\lambda_0}\right] + C\mu_i(p). \tag{24}$$

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

$$(\lambda_1+\lambda_2)p(1-p)v_i' + (r+(\lambda_1+\lambda_2)p)v_i = \big(\lambda_i\alpha + \lambda_j(1-\alpha)\big)\,p. \tag{25}$$

The solution to the above differential equation is

$$v_i^{rr}(p) = \frac{\lambda_i\alpha+\lambda_j(1-\alpha)}{r+\lambda_1+\lambda_2}p + C\mu(p). \tag{26}$$

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 $p \le p_1^*$, we have $\lambda_1 p[1-2v-2v'(1-p)] \le \lambda_0(1-2v)$ 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 cartel'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 $p \le p_1^*$, this strategy also constitutes the best response of firm i. Consider $p \le p_1^*$. 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

$p = p_1^*$, $v_i = \frac{2\alpha\lambda_0+\lambda_1 p(1-2\alpha)}{r+2\lambda_0}$. Since $\alpha \geq \frac{1}{2}$, $\frac{2\alpha\lambda_0+\lambda_1 p(1-2\alpha)}{r+2\lambda_0}$ is non-increasing in p. This implies that, for all $p > p_1^*$, we have $v_i > \frac{2\alpha\lambda_0+\lambda_1 p(1-2\alpha)}{r+2\lambda_0}$.

To show uniqueness, consider again the range $p \leq p_1^*$ and suppose that a firm adopts 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 (10) and (12), it follows immediately from $\hat{p} < p_1^*$ that $v_i < \frac{\lambda_0}{r+2\lambda_0} \leq \frac{2\alpha\lambda_0+\lambda_1 p(1-2\alpha)}{r+2\lambda_0}$ 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 $\mathscr{N} \cap (p_1^*, \check{p})$. 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 $v_j'(\tilde{p}-) \leq 0$. (10) and (12) imply that we must have $(k_i, k_j)(\tilde{p}) = (1,0)$. The Bellman equation (1) immediately implies 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 (16) and (18), 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 $p \in (p_1^*, p_2^*)$, (16) 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, (16) 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 (18), 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 Theorem 1

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 $\lambda_1 > \lambda_0$, and $\alpha > \frac{1}{2} > \frac{\lambda_0}{r+2\lambda_0}$. Direct computation shows $v_1^{rs'}(\frac{\lambda_0}{\lambda_1}+) = 0$.

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 $v_2^{rs'}(\frac{\lambda_0}{\lambda_1}+) = 0$, $C_2^{rs} < 0$ if $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, $C_2^{rs} > 0$ if $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$, and $C_2^{rs} = 0$ if $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$. Using the ODEs (23) and (25), together with value matching and the definition of $\hat{p}_2$,[9] establishes smooth pasting at $\hat{p}_2$. Thus, v_2 is continuously differentiable. On $(\hat{p}_2, 1)$, it is strictly decreasing and concave on $(p_1^*, \hat{p}_2)$ if $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, strictly increasing and convex if $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$ and flat at $\frac{\lambda_0}{r+2\lambda_0}$ if $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$. As we shall show below, it is convex on $(\hat{p}_2, 1)$.

We now show that $\hat{p}_2(\alpha)$ is well-defined, i.e. that there exists a unique $\hat{p}_2(\alpha) \in (p_1^*, 1)$ such that $F(\hat{p}_2(\alpha), \alpha) = 0$, where the differentiable function F is defined by

$$F(p, \alpha) = v_2^{rs}(p) - \frac{\lambda_0 \alpha(\lambda_1 + \lambda_2) + \lambda_1 \lambda_2 p(1-2\alpha)}{r\lambda_2 + \lambda_0(\lambda_1 + \lambda_2)}.$$

At $p = p_1^*$, $v_2^{rs}(p) = \frac{\lambda_0}{r+2\lambda_0}$ and $\frac{\lambda_0 \alpha(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p_1^*(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} = \frac{\alpha\lambda_0\lambda_1+(1-\alpha)\lambda_0\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Thus, we have

$$\frac{\lambda_0 \alpha(\lambda_1 + \lambda_2) + \lambda_1 \lambda_2 p(1-2\alpha)}{r\lambda_2 + \lambda_0(\lambda_1 + \lambda_2)} - v_2^{rs}(p) = \frac{\lambda_0(\alpha r + (2\alpha - 1)\lambda_0)(\lambda_1 - \lambda_2)}{(r\lambda_2 + \lambda_0(\lambda_1 + \lambda_2))(r + 2\lambda_0)} > 0$$

as $\alpha \geq \frac{1}{2}$. Thus, $F(p_1^*, \alpha) < 0$ for all $\alpha \geq \frac{1}{2}$.

At $p = 1$, we have $v_2^{rs}(p) = \frac{\alpha\lambda_0+(1-\alpha)\lambda_1}{r+\lambda_0+\lambda_1}$ and $\frac{\lambda_0\alpha(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} = \frac{\alpha\lambda_0(\lambda_1+\lambda_2)+(1-2\alpha)\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Let $A = \frac{\alpha\lambda_0+(1-\alpha)\lambda_1}{r+\lambda_0+\lambda_1} - \frac{\alpha\lambda_0(\lambda_1+\lambda_2)+(1-2\alpha)\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Direct computation shows that A is strictly increasing in α, and at $\alpha = \frac{1}{2}$, $A > 0$. Thus, for all $\alpha \geq \frac{1}{2}$, we have $F(1, \alpha) > 0$.

If $\alpha < \frac{r+\lambda_0}{r+2\lambda_2}$, v_2^{rs} is strictly increasing, while $p \mapsto \frac{\lambda_0\alpha(\lambda_1+\lambda_2)+(1-2\alpha)\lambda_1\lambda_2}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ is decreasing. Thus, we can conclude that there exists a unique $\hat{p}_2(\alpha) \in (p_1^*, 1)$ such that $F(\hat{p}_2(\alpha), \alpha) = 0$.

If $\alpha \geq \frac{r+\lambda_0}{r+2\lambda_2}$, both v_2^{rs} and $p \mapsto \frac{\lambda_0\alpha(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ are (weakly) decreasing in p. The slope of v_2^{rs} is bounded below by $\frac{\lambda_1}{r+\lambda_0+\lambda_1}(1 - \frac{r+2\lambda_0}{r+\lambda_0}\alpha)$, while the slope of $\frac{\lambda_0\alpha(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ is $\frac{\lambda_1\lambda_2(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. Since

$$B = \frac{\lambda_1\lambda_2(2\alpha - 1)}{r\lambda_2 + \lambda_0(\lambda_1 + \lambda_2)} - \frac{\lambda_1}{r + \lambda_0 + \lambda_1}(\frac{r + 2\lambda_0}{r + \lambda_0}\alpha - 1)$$

is strictly increasing in α and at $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$ we have $B > 0$, we can conclude that, for all $\alpha \geq \frac{r+\lambda_0}{r+2\lambda_0}$, we have $B > 0$. Thus, the slope of $p \mapsto \frac{\lambda_0\alpha(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ is strictly lower than the lower bound on the slope of v_2^{rs}, so that we can conclude that there exists a unique $\hat{p}_2(\alpha) \in (p_1^*, 1)$ such that $F(\hat{p}_2(\alpha), \alpha) = 0$.

That $\hat{p}_2(\frac{1}{2}) = p_2^*$ follows immediately from the defining equations. By player 2's Bellman equation (4), smooth pasting at $\hat{p}_2$ implies that $\lambda_2 \hat{p}_2(\alpha - v_2(\hat{p}_2) - (1-\hat{p}_2)v_2'(\hat{p}_2)) = \lambda_0(\alpha - v_2(\hat{p}_2))$. As $v_2'(\hat{p}_2) < 0$, $v_2'(\hat{p}_2) = 0$ and $v_2'(\hat{p}_2) > 0$ in the cases $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$ and $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$, respectively, this implies $\hat{p}_2 < \frac{\lambda_0}{\lambda_2}$ if $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, $\hat{p}_2 = \frac{\lambda_0}{\lambda_2}$ if $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$ and $\hat{p}_2 > \frac{\lambda_0}{\lambda_2}$ if $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$.

We shall now show that the cutoff $\hat{p}_2$ is strictly decreasing in α. Direct computation shows that $\frac{\partial F}{\partial p} > 0$, so that the sign of $\frac{d\hat{p}_2}{d\alpha}$ is the opposite of the sign of $\frac{\partial F}{\partial \alpha}$. Direct computation shows that $\frac{\partial F}{\partial \alpha}(p, \alpha)$ is independent of α and a strictly increasing continuous function of p, which is strictly

[9] We omit the argument of $\hat{p}_2(\alpha)$ whenever convenient to do so.

negative at $p=\frac{\lambda_0}{\lambda_1}$ and strictly positive at $p=\frac{\lambda_0}{\lambda_2}$. Thus, there exists a unique $\tilde{p}\in(\frac{\lambda_0}{\lambda_1},\frac{\lambda_0}{\lambda_2})$ such that $\frac{\partial F}{\partial\alpha}$ switches its sign from negative to positive as p increases to $\tilde{p}$.

Since $\hat{p}_2\geq\frac{\lambda_0}{\lambda_2}$ for $\alpha\in\left[\frac{1}{2},\frac{r+\lambda_0}{r+2\lambda_0}\right]$, it follows that $\hat{p}_2$ is strictly decreasing in α in this range. Now suppose there exists an $\hat{\alpha}\in(\frac{r+\lambda_0}{r+2\lambda_0},1]$ such that $\hat{p}_2(\hat{\alpha})=\tilde{p}$. Then, $\frac{\partial F}{\partial\alpha}(\hat{p}_2(\hat{\alpha}),\hat{\alpha})=0$. As $\hat{\alpha}>\frac{r+\lambda_0}{r+2\lambda_0}$, $\tilde{p}<\frac{\lambda_0}{\lambda_2}$. Since all higher derivatives of this function of α are also 0 at $\hat{\alpha}$, it follows that $\hat{p}_2(\frac{1}{2})=\hat{p}_2(\hat{\alpha})=\tilde{p}$. As, by Proposition 3, $\hat{p}_2(\frac{1}{2})=p_2^*>\frac{\lambda_0}{\lambda_2}$, we get the following chain of inequalities: $\frac{\lambda_0}{\lambda_2}<\hat{p}_2(\frac{1}{2})=\hat{p}_2(\hat{\alpha})=\tilde{p}<\frac{\lambda_0}{\lambda_2}$, a contradiction.

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 $v_2(p)\leq\frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ for all $p\in(\frac{\lambda_0}{\lambda_1},\hat{p}_2]$ by construction, $k_2=0$ satisfies the Bellman equation as well.

Finally, we consider the range of beliefs $(\hat{p}_2,1]$, and first establish strict convexity of v_2 in this range. To do so, we consider the function $\tilde{v}_2(p)=\frac{\lambda_2\alpha+\lambda_1(1-\alpha)}{r+\lambda_1+\lambda_2}p+\tilde{C}_2\mu(p)$, where the constant $\tilde{C}_2$ is implicitly defined by $\tilde{v}_2(p_1^*)=\frac{\lambda_0}{r+2\lambda_0}$. This immediately implies that $\tilde{C}_2>0$. By our previous step, player 2 uniquely best-responds by playing safe in the range $(p_1^*,\hat{p}_2)$, which implies that $v_2>\tilde{v}_2$ on $(p_1^*,\hat{p}_2)$. Therefore, $v_2(\hat{p}_2)=\frac{\lambda_2\alpha+\lambda_1(1-\alpha)}{r+\lambda_1+\lambda_2}\hat{p}_2+C_2^{rr}\mu(\hat{p}_2)\geq\tilde{v}_2(\hat{p}_2)$, Thus, $C_2^{rr}>0$, and v_2 is strictly convex on $(\hat{p}_2,1)$.

By convexity of v_2, smooth pasting at $\hat{p}_2$ and the fact that the graph of v_2^{rs} intersects the graph of $p\mapsto\frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2(1-2\alpha)p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ from below , $v_2'>\frac{\lambda_1\lambda_2(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ in the range $(\hat{p}_2,1)$. This implies that $v_2(p)>\frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2(1-2\alpha)p}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ for all p in this range, and hence player 2 is playing a best response at these beliefs as well.

To show the best-response property on $(\hat{p}_2,1)$ for player 1 as well, we consider the function $\tilde{v}_1(p)=\frac{\lambda_1\alpha+\lambda_2(1-\alpha)}{r+\lambda_1+\lambda_2}p+\tilde{C}_1\mu(p)$, where the constant $\tilde{C}_1$ is implicitly defined by $\tilde{v}_1(\hat{p}_2)=\frac{\lambda_0}{r+2\lambda_0}$. From (25), it follows that, at any belief $\tilde{p}$ such that $\tilde{v}_1(\tilde{p})=\frac{\lambda_0}{r+2\lambda_0}$, we have $\tilde{v}_1{}'(\tilde{p})>0$ if and only if $\tilde{p}>\frac{r\lambda_0}{\lambda_0(2\alpha-1)(\lambda_1-\lambda_2)+r(\lambda_1\alpha+\lambda_2(1-\alpha))}$. We will now distinguish the cases (1.) $\alpha\geq\frac{r+\lambda_0}{r+2\lambda_0}$ and (2.) $\alpha<\frac{r+\lambda_0}{r+2\lambda_0}$. Direct computation shows that $\frac{r\lambda_0}{\lambda_0(2\alpha-1)(\lambda_1-\lambda_2)+r(\lambda_1\alpha+\lambda_2(1-\alpha))}\leq\frac{\lambda_0}{\lambda_1}$ if and only if $\alpha\geq\frac{r+\lambda_0}{r+2\lambda_0}$. As $\hat{p}_2>\frac{\lambda_0}{\lambda_1}$, we can conclude that, in case (1.), $\tilde{v}_1>\frac{\lambda_0}{r+2\lambda_0}$ 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 $p\in(\hat{p}_2,1)$. Now, let us turn to case (2.). Direct computation shows that $\frac{\lambda_0}{\lambda_2}>\frac{r\lambda_0}{\lambda_0(2\alpha-1)(\lambda_1-\lambda_2)+r(\lambda_1\alpha+\lambda_2(1-\alpha))}$. Since $\hat{p}_2>\frac{\lambda_0}{\lambda_2}$ in case (2.), we can infer that $\tilde{v}_1(p)>\frac{\lambda_0}{r+2\lambda_0}$ 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 again conclude that $v_1^{rr}(p)>\tilde{v}_1(p)$ for $p\in(\hat{p}_2,1)$. The fact that, for $p>p_1^*$, $\frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_1+\lambda_0(\lambda_1+\lambda_2)}<\frac{\lambda_0}{r+2\lambda_0}$ implies that firm 1 is playing a best response on $(\hat{p}_2,1)$ in case (2.) as well.

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 $\frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} > \frac{\lambda_0[\alpha\lambda_1+(1-\alpha)\lambda_2]}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} > \frac{\lambda_0}{r+2\lambda_0} > 0$. 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

$$\frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)} > v_2(p)$$

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 $p_l < \frac{\lambda_0}{\lambda_1} < \frac{\lambda_0}{\lambda_2}$, (21) 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, $v_2 < \frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$ 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 (21), 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 $\frac{\lambda_0}{\lambda_1}$.

For the range $(p_1^*, \hat{p}_2]$, we shall distinguish two cases: (1.) $\alpha \geq \frac{r+\lambda_0}{r+2\lambda_0}$ and (2.) $\alpha < \frac{r+\lambda_0}{r+2\lambda_0}$. We start with case (1.), and shall argue next 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, if $\alpha > \frac{r+\lambda_0}{r+2\lambda_0}$, the payoff function of firm 2 (22) 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. If $\alpha = \frac{r+\lambda_0}{r+2\lambda_0}$, $p^{'} < \hat{p}_2 = \frac{\lambda_0}{\lambda_2}$ implies that the payoff function of firm 2 (22) drops below $\frac{\lambda_0}{r+2\lambda_0}$ to the immediate right of $p^{'}$, so that firm 2 is not playing a best response in some right-neighbourhood of $p^{'}$.

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 $p_l^{''}$. We have already established that, in any equilibrium, either $(k_1,k_2) = (1,0)$ or $(k_1,k_2) = (1,1)$ prevails throughout $(p_1^*, p_l^{''}]$. Using the ODEs (21) and (25) and the assumption $\alpha \geq \frac{r+\lambda_0}{r+2\lambda_0}$, one can show that firm 1's payoff satisfies $v_1(p_l^{''}-) > \frac{\lambda_0}{r+2\lambda_0}$, 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 $v_2(p_l^{'''}) = v_2^{rs}(p_l^{'''}) < \frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$, where the inequality follows from $p_l''' < \hat{p}_2$, implying that firm 2 has a profitable deviation.

Now, let us turn to case (2.) and suppose there exists an equilibrium with the feature that there exists a $p \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2)$ such that $(k_1, k_2) \neq (1,0)$, and let $p_l = \inf\{p \in (\frac{\lambda_0}{\lambda_1}, \hat{p}_2) : (k_1, k_2) \neq (1,0)\}$. Thus, $v_1(p_l) = v_1^{rs}(p_l) > \frac{\lambda_0}{r+2\lambda_0} > \frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_1+\lambda_0(\lambda_1+\lambda_2)}$ and $v_2(p_l) = v_2^{rs}(p_l) < \frac{\alpha\lambda_0(\lambda_1+\lambda_2)+\lambda_1\lambda_2 p(1-2\alpha)}{r\lambda_2+\lambda_0(\lambda_1+\lambda_2)}$. This implies that $k_1 = 1$ is a strictly dominant action for firm 1 in some right-neighbourhood of p_l, while $k_2 = 0$ is firm 2's unique best response in this range, a contradiction. 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 shall 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 $\tilde{p}$. Our verification arguments imply that firm 1 is not playing a best response at beliefs just above $\tilde{p}$. A similar argument to above furthermore establishes that firm 1 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 our equilibrium construction, it follows that the unique best response of firm 2 is to choose R, which concludes the proof.