WARWICK
ECONOMICS

CRETA
Centre for Research in Economic Theory and its Applications

Discussion Paper Series

Experimentation in Dynamic R&D Competition

Anastasios Dosis & Abhinay Muthoo

(This paper also appears as
Warwick Economics Research Papers series No: 1214)

July 2019 No: 52

CRETA
Centre for Research in Economic Theory and its Applications
Department of Economics
University of Warwick, Coventry,
CV4 7AL, United Kingdom

warwick.ac.uk/fac/soc/economics/research/centres/creta/papers

Experimentation in Dynamic R&D Competition*

Anastasios Dosis[†] Abhinay Muthoo[‡]

February 4, 2019

Abstract

We study a two-stage, winner-takes-all, R&D race, in which, at the outset, firms are uncertain regarding the viability of the project. Learning through experimentation introduces a bilateral (dynamic) feedback mechanism. For relatively low-value products, the equilibrium stopping time coincides with the socially efficient stopping time although firms might experiment excessively in equilibrium; for relatively high-value products, firms might reduce experimentation and stop rather prematurely due to the fundamental free-riding effect. Perhaps surprisingly, a decrease in the value of the product can spur experimentation.

KEYWORDS: Experimentation, learning, dynamic R&D competition, inefficiency

JEL CLASSIFICATION: C73, D83, O31, O32

1 Introduction

Research and Development (R&D) exhibits at least three characteristics. First, it is a dynamic, risky process. This is well documented in the pharmaceutical industry where the development of a new drug requires several phases to be completed.[1] It has been esti-

*We are grateful to Alex Gershkov and Motty Perry for many insightful comments. We also thank Malin Arve, Pauli Murto, Herakles Polemarchakis, Petros Sekeris and Juuso Valimaki as well as participants at the 11th Conference on Research on Economic Theory and Econometrics (Milos), the 21st European Workshop on General Equilibrium Theory (Exeter), the Sixth International Conference on Game Theory and Management (St. Petersburg), the 4th World Congress of the Game Theory Society (Istanbul), the Conference on Auctions, competition, regulation, and public policy (Lancaster Management School), the ESSEC Economics Brown Bag seminar and the THEMA theory seminar for helpful comments and remarks. All errors are our own.

[†]Department of Economics - ESSEC Business School and THEMA, 3 Av. Bernard Hirsch, B.P. – 50105, Cergy, 95021, France, Email: dosis@essec.com.

[‡]Department of Economics - University of Warwick, Coventry CV4 7AL, United Kingdom, Email: A.Muthoo@warwick.ac.uk

[1]Research usually begins in the lab where researchers identify promising compounds for development. If a compound passes this stage, it moves to pre-clinical (tests in animals) and clinical (tests in humans) research that are prerequisites for FDA approval. See https://www.fda.gov/ForPatients/Approvals/Drugs/default.htm.

mated that less than 1% of the compounds examined in pre-clinical trials make it into clinical trials and only 20% of the compounds entering clinical trials survive the development process and gain FDA approval.[2] The pre-clinical and clinical testing phases generally take more than a decade to complete. Second, firms (or research teams) usually compete with one another for being the first to cross the finish line. Winners are usually allocated a patent which provides exclusive rights to market the product and/or entertain a first-mover advantage. Last, given its dynamic nature, R&D entails learning as firms continuously update their beliefs regarding the viability of the project. As is well known, information (i.e., learning) has the characteristics of a public good; a discovery by one firm has a positive impact on other firms that experiment on a similar project.

In this paper, we study a model of dynamic R&D competition that encompasses all the characteristics described above. Two firms engage in a two-stage winner-takes-all race. The first firm that completes the two stages first is allocated a patent and monopoly profits for a given period. In the example given above on the development of a new drug, the research stage may refer to research in the lab whilst the development stage may refer to pre-clinical and clinical trials. Both firms start the race equipped with similar bandit technologies on which they can experiment. The technologies are perfectly correlated; they are either both armed in good or in a bad state. A technology that is armed in a good state generates a success with positive probability; a technology that is armed in a bad state never generates a success. This implies that with some positive probability a stage is never completed successfully as is evident, for instance, in the pharmaceutical industry. Firms independently and simultaneously decide on the intensity of experimentation with every unit spent on experimentation having an opportunity cost that arises from exploiting an existing certain project. Crucially for our analysis, firms decide on their intensity of experimentation on a period by period basis; a firm might stop experimenting for some time but restart at any point in the future.

Uncertainty is gradually resolved as firms experiment on a given technology. Every failure is a signal that the technology is armed in a bad state; hence, it makes firms more pessimistic regarding its viability to generate the required breakthrough. An important assumption is that firms observe one another's actions and outcomes which introduces a bilateral (dynamic) feedback mechanism. On the one hand, a failure of one firm imposes a *negative information externality* to the rival as is now more likely that the technology is armed in a bad state; a success imposes a *positive information externality* as it now becomes commonly known that advancement to the next stage is feasible. On the other hand, success by one firm imposes a *negative competition externality* on the rival as the latter needs to catch up in the race. In other words, the rival has a lag in the race arising from the fact that it still needs to invest resources and successfully pass the research stage before competing head-to-head with the leader. Our goal is a detailed study of the interplay between these two countervailing externalities.

We characterise the unique symmetric subgame perfect Nash equilibrium of the game

[2] Discontinuation of research can result from toxicity, carcinogenicity or any other manufacturing difficulties that researchers discover during pre-clinical or clinical trials. See DiMasi (1995), DiMasi, Hansen, and Grabowski (2003) and DiMasi, Grabowski, and Hansen (2016).

by partitioning the set of parameters into two sets. For relatively low-value products, advancement of one firm to the development stage discourages the rival from remaining in the race; the negative competition externality offsets the positive information externality. This result is in line with early papers in the patent races literature (e.g., Fudenberg et al. (1983), Harris and Vickers (1985)), who argue that an early advantage is sufficient for the race to degenerate into a monopoly.

Perhaps the most interesting case arises for relatively high-value products. In this case, the positive information externality offsets the negative competition externality; hence, the laggard finds it profitable to stay in the race. We show that at the *switching point* (i.e., the cutoff value above which the positive externality offsets the negative externality), there is a downward jump discontinuity in the payoff of the leader. Intuitively, the leader would prefer the race to degenerate into a monopoly instead of competing with the laggard even if the former enjoys a first-mover advantage.

We characterise the equilibrium stopping time as a function of the value of the product. We find that the stopping time is everywhere continuously decreasing in the value of the product but exhibits an upward jump discontinuity at the switching point. This discontinuity arises from the discontinuity in the continuation payoff of the leader; advancement by one firm to the development stage is accompanied by an increase in experimentation by the rival which has a negative impact on the continuation payoff of the firm that advanced first. The model implies that (i) the intensity of experimentation decreases in the belief, and, (ii) around the discontinuity, a decrease in the product's value can spur experimentation.

We study the welfare properties of the symmetric equilibrium. The symmetric equilibrium is everywhere inefficient: when at least one firm has advanced to the development stage, firms experiment excessively because they do not internalise the negative externalities they impose on one another (e.g., Loury (1979), Lee and Wilde (1980)); when both firms are in the research stage, firms might either experiment excessively or insufficiently. The latter is a result of the free-riding effect and is particularly prevalent in the strategic experimenation literature (e.g., Keller, Rady, and Cripps (2005)).

Casual observation suggests that the dynamic feedback mechanism can be important in real-world races. When Bednorz and Müller (1986) announced that they had discovered superconductivity in ceramic materials at 35^o Kelvin - a temperature that seemed impossible before - physicists realised that other high-temperature superconducting material are feasible. Following a conference that took place in New York, during which Wu et al. (1987) presented superconducting materials at much higher temperature (around 90^o Kelvin), "*...hundreds of scientists, Nobel laureates as well as the greenest graduate students, went on a work rampage. It was indeed a race, a race to obtain fame and patents.*[3] As it is evident in this example, discovery by one firm might result in an increase of effort by rivals to be the first to cross the line.

□ **Related Literature**. The paper is related to several strands in the literature. Early papers by Loury (1979), Lee and Wilde (1980) and Dasgupta and Stiglitz (1980) analyse

[3] Vidali (1993).

(effectively static) patent races in which firms invest in R&D only at the outset and investment determines the speed of innovation. All papers show that R&D investment is socially excessive. We also find that R&D investment is excessive but this is more prevalent either in more advanced stages or for relatively low-value products.

Reinganum (1981, 1982), Fudenberg et al. (1983) and Harris and Vickers (1985, 1987) allow for dynamic models in which firms can only increase their investments (i.e., once a firm drops out, it can never re-enter the race). In Reinganum (1981, 1982), firms accumulate knowledge; when knowledge has the characteristics of a private good (i.e. it can only be used by the firm that accumulates it), then R&D investment is excessive; when knowledge has the characteristics of a public good, then investment is insufficient. Fudenberg et al. (1983) analyse a multistage race and provide conditions under which a race will be characterised by vigorous competition or will degenerate into monopoly. Key is whether the laggard has time to "leapfrog". Harris and Vickers (1987) show that the leader exerts higher effort than the laggard, and effort increases as the gap between competitors decreases. A similar result in obtained in Grossman and Shapiro (1987) who extend the model of Lee and Wilde (1980) to two stages. They find that (i) when both firms are in the development stage, competition is the most intense; (ii) the leader invests more than the laggard. Although in our model the investment decision is binary (i.e., invest or not invest), we also find that for relatively low-value products, the leader stays active whilst the follower drops out.

Closely related are the papers by Choi (1991), Malueg and Tsutsui (1997), Mason and Valimaki (2010) and Moscarini and Squintani (2010) who also study related model of dynamic R&D competition. Similarly to our paper, Choi (1991) studies a model with hazard rate uncertainty and identifies the different externalities that might affect R&D investment. Unlike our paper, he assumes that once firms exit the race, they are unable to re-enter.[4] In our model firms are free to exit and re-enter anytime. Re-entry is the underlying mechanism of the discontinuity in the stopping times. Unlike Choi (1991), we proceed to a welfare analysis and discuss asymmetric equilibria. Malueg and Tsutsui (1997) study a one stage race with hazard rate uncertainty. One of their main results is that as firms become more pessimistic, they reduce investment and possibly exit the race. In our model the interesting insights arise only due to the two stages.

This paper draws insights from the strategic experimentation literature. Bolton and Harris (1999) study two-armed bandit models with several agents and Keller, Rady, and Cripps (2005) extend their framework to exponential bandits. Building on Keller, Rady, and Cripps (2005), Bimpikis, Ehsani, and Mostagir (2015) and Halac, Kartik, and Liu (2017) study dynamic contest design in which the planner commits to a disclosure policy and reward scheme. Our paper takes the R&D game as given and analyses its equilibria.[5]

[4] In p. 608, Choi (1991) writes: "*... But in models of hazard rate uncertainty, the payoff structure of the race could be that of a waiting game. Every firm might prefer the other firms to enter the race first. After observing the performance of other firms, it will enter if the other firms do well and the race is considered to exhibit a high enough hazard rate. Another possibility is that both firms enter the race, and after some time without success by either firm, one of them drops out. The firm that drops out could reenter once the remaining firm succeeds in the intermediate-stage discovery...*"

[5] Related are the paper by Kremer, Mansour, and Perry (2014) Che and Horner (2015), Heidhues, Rady,

The remainder of the paper is organised as follows. In Section 2, we present the model. In Section 3, we analyse the unique symmetric subgame perfect Nash equilibrium. In Section 4, we study the welfare properties of the unique symmetric equilibrium. In Section 5 we discuss the relevance of the model, the robustness of our results as well as further potential extensions. Section 7 concludes the paper. All formal proofs are provided in the appendix.

2 THE MODEL

There is a potentially infinite number of periods. Two symmetric, risk-neutral firms are engaged in a race to launch a new product (e.g., a new drug). Development of the new product requires a firm to successfully complete two stages: *research* and *development*. Let s denote the stage, where $s = r, d$. The first firm that successfully completes both stages receives a patent and monopoly profits equal to v.[6] Advancement to the development stage requires a firm to successfully complete the research stage, which implies that the two stages are consecutive steps in the production process. Although a firm that completes the research stage qualifies to the development stage, the rival could potentially qualify to the development stage by also successfully completing the research stage. We call a firm that qualified first to the development stage the *leader* and the firm that is still at the research stage the *laggard*. In case both firms complete the two stages simultaneously, then the patent is allocated to one of them with equal probability.[7] For simplicity and to highlight the potential inefficiencies, we assume that firms do not discount the future. However, all our results remain valid even under discounting. We discuss this point at the end of the paper.

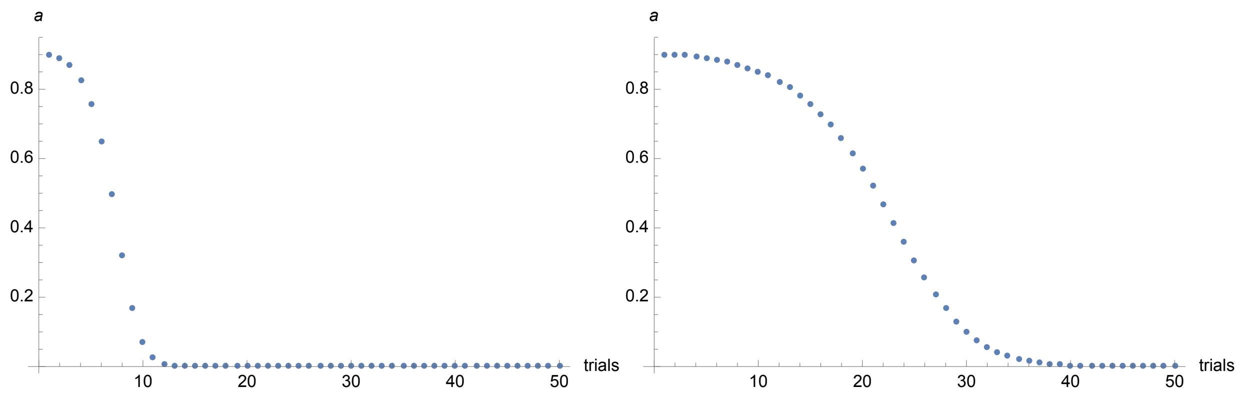
In every stage, firms each have symmetric bandit technologies on which they can experiment. Firms choose their intensity of experimentation $x \in [0, 1]$. Every dollar spent on experimentaion has an opportunity cost c. This cost can be interpreted as the the profit that the firm can earn by operating an alternative project of certain returns. An important assumption is that each firm decides on its intensity of experimention on a period by period basis. In particular, we assume that a firm that has stopped experimenting for some time can freely re-commence at any future point.

Every technology is armed either in a good or a bad state. As is standard in the strategic experimentation literature, the two technologies are perfectly correlated; either both are in a good or a bad state. The common prior belief that the technology is in a good state is $\alpha_0 < 1$. A technology generates a *success* or a *failure*. One success is required for a firm to advance from the research to the development stage and two successes are required for a firm to be allocated the patent. We assume that successes occur independently in the two technologies. If a technology is armed in a good state and a firm chooses an intensity x, then the probability of a firm generating a success is px, where $p < 1$; if a technology

and Strack (2015) and Bimpikis and Drakopoulos (2018), who analyse how information disclosure can improve learning.

[6] This corresponds to the sum of the profits from the effective life of the patent.

[7] One can consider different patent allocation rules without substantially change the results. This allocation rule is chosen for simplicity.



(a) $a_0 = 0.9$, $p = 0.1$, $x = 1$ (b) $a_0 = 0.9$, $p = 0.01$,$x = 1$

Figure 1

is armed in a bad state the correspondng probability of generating a success is zero regardless of the intensity of experimentation.[8] Therefore, only if the technology is armed in a good state, success is ever possible; if the technology is armed in a bad state, success is never possible. The two states represent the potential uncertainty that firms face when undertaking R&D; firms do not know if the product can ever be developed.

Uncertainty is resolved gradually if firms experiment; if at least one firm succeeds in the research stage, both firms become certain that the technology is in a good state; a failure delivers bad news and both firms become more pessimistic about the viability of the technology to generate the breakthrough. In particular, suppose that the belief is α and only one firm experiments at intensity x, then, following a failure, the belief that the technology is in a good state is downgraded to

$$\alpha' = \frac{\alpha(1-px)}{(1-\alpha)+\alpha(1-px)} \tag{1}$$

The posterior belief deteriorates faster the higher is p and/or x; i.e., the higher is the probability of success if the technology is in a good state, the more informative is a failure. Figure 1 represents two cases for $\alpha_0 = 0.9$. In Panel 1a, $p = 0.1$, $x = 1$ and in Panel 1b, $p = 0.01$, $x = 1$. One can see that the posterior belief deteriorates considerably faster in Panel 1a.

3 The Unique Symmetric Equilibrium

■ **Symmetric Equilibrium and Notation.** We assume that firms each simultaneously and independently take actions every period by having observed one another's actions and outcomes in the previous periods; therefore, we study a dynamic game of complete information. We characterise the unique symmetric subgame perfect Nash equilibrium.

[8]We assume that the intensity of experimentation cannot exceed an upper bound; hence, there is always an upper bound in the probability a firm can generate a success. Because experimentation is costly, firms will never choose an intensity greater than one.

v	Value of the product
p	Marginal probability of success if the state is good
c	Marginal cost of experimentation
α_0	Prior that the state is good
x^i	Firm i's intensity of experimentation
$\bar{x}_{ss'}$	Symmetric equilibrium intensity of experimentation of the firm that is in stage s, conditional on the rival being in stage s', where $ss' \in \{rd, dr, dd\}$
$\bar{x}_{rr}(\alpha)$	Symmetric equilibrium intensity of experimentation when both firms are in research and the belief is α
$\bar{V}_{ss'}$	Symmetric equilibrium continuation payoff of a firm that is in stage s conditional on the rival being in stage s', where $ss' \in \{rd, dr, dd\}$

Table 1: Notation

A strategy for each firm defines the intensity of experimentation every period for any history of play. To simplify notation, let $\bar{x}_{ss'}$ and $\bar{V}_{ss'}$ denote the equilibrium intensity of experimentation in stage s, conditional on the other firm being in stage s', and the equilibrium continuation payoff of a firm that is in stage s if the rival firm is in stage s', where $ss' \in \{rd, dr, dd\}$. By definition, in a symmetric equilibrium in which both firms are in the same stage, the intensities of experimentation and continuation payoffs of the two firms are equal. Note also that if at least one firm qualifies to the development stage, the problem becomes stationary; hence, if a decision is optimal in one period, it remains optimal every period. Let also $\bar{x}_{rr}(\alpha)$ denote the intensity of experimentation conditional on both firms being in the research stage and the posterior belief is α.

□ **Both Firms in the Development Stage**. Suppose first that both firms have qualified to the development stage. We impose the following assumption which implies that experimentation is always profitable for a sufficiently high belief.

Assumption 1. $v \geq \frac{2c}{p}$

The equillibrium intensity of experimentation depends on the relative value of the product. Given Assumption 1, it is relatively straightforward to see that firms each choose a strictly positive intensity. Suppose that in every period, firm i chooses (constant) intensity x^i and firm j chooses (constant) intensity x^j; then, the (static) expected payoff of firm i is

$$\Big\{\frac{p^2 x^i x^j}{2} + p x^i(1 - p x^j)\Big\} v - c x^i = \Big\{p(1 - \frac{p x^j}{2}) v - c\Big\} x^i \tag{2}$$

Due to Assumption 1, it is clear that Eq. (2) is strictly increasing in x^i for every x^j. This implies that in the unique symmetric equilibrium, both firms will choose to experiment at a maximum intensity in every period. We summarise the analysis in the following lemma.

Lemma 1. *Suppose that both firms have qualified to the development stage; then, in the unique symmetric equilibrium* $\bar{x}_{dd} = 1$ *and*

$$\bar{V}_{dd} = \frac{1}{2}\Big(v - \bar{v}_{dd}\Big) \tag{3}$$

where

$$\bar{v}_{dd} = \frac{2c}{p(2-p)} \tag{4}$$

Proof. See the Appendix. □

It is insightful to consider the effect of changes in parameters on the continuation value. It is straighforward from inspection of Eq. (4) that $\bar{v}_{dd}$ is strictly increasing in c and p; hence, by inspection of Eq. (3), one can straightforwardly deduce that (ceteris paribus) an increase in c or an increase in p decrease $\bar{V}_{dd}$. The effect of p on $\bar{v}_{dd}$ and $\bar{V}_{dd}$ is a consequence of a *strategic effect*: an increase in the probability of success of both firms makes experimentation less attractive for each firm conditional on the rival experimenting at a maximum intensity.

□ **One Firm in the Research and One in the Development Stage**. Having characterised the equilibrium behaviour of the two firms if both have advanced to the development stage and the corresponding equilibrium continuation payoffs, we can now study the subgame in which one firm is in the research stage and the rival is in the development stage.

It is relatively straightforward that the leader exploits its position and, given Assumption 1, always experiments at a maximum intensity. For relatively low-value products, the leader remains a monopolist; for relatively high-value products, the laggard stays in the race and also experiments at a maximum intensity.[9]

Given stationarity, the necessary and sufficient condition for the laggard to be active is

$$p(1-p)\bar{V}_{dd} - c \geq 0 \tag{5}$$

where $p(1-p)$ represents the probability that the laggard succeeds and the leader fails (conditionally on both experimenting at a maximum intensity) in which case both firms find themselves in the development stage and each earns a continuation payoff equal to $\bar{V}_{dd}$.[10] Straighforward algebra reveals that the laggard remains active if and only if the value of the product is greater than

$$\bar{v}_{rd} \equiv \bar{v}_{dd} + \frac{2c}{p(1-p)} \tag{6}$$

[9]Conventionally, we assume that, conditionally on the rival's strategy, when a firm is indifferent between staying active or not it remains active.

[10]The (static) expected payoff of the laggard if it experiments at intensity x_{rd} whilst the leader experiments at intensity x_{dr} is:

$$\Big\{p(1-px_{dr})\bar{V}_{dd} - c\Big\}x_{rd}$$

Because, as we argued above, the leader always experiments at maximum intensity $x_{dr} = 1$; hence, the necessary and sufficient condition for the laggard to experiment is given in Eq. (5).

Eq. (6) implies that, conditionally on the leader experimenting at a maximum intensity, a markup over $\bar{v}_{dd}$ is required to incentivise the laggard to remain active given that this needs to catch up in the race. From inspection of Eq. (6), one can see that $\bar{v}_{rd}$ is strictly increasing in c and p – i.e., an increase in the cost of experimentation or an increase in the probability the breakthrough occurs both decrease the threshold value of the product above which the laggard experiments.

We summarise the unique symmetric equilibrium in the following lemma.

Lemma 2. *Suppose that one firm is in the research stage and the other firm is in the development stage; then, in the unique symmetric equilibrium* $\bar{x}_{dr} = 1$ *and*

$$\bar{x}_{rd} = \begin{cases} 0, & \textit{if} \quad v \leq \bar{v}_{rd} \\ 1, & \textit{otherwise} \end{cases} \tag{7}$$

The equilibrium continuation payoffs of the two firms are respectively

$$\bar{V}_{rd} = \begin{cases} 0, & \textit{if} \quad v \leq \bar{v}_{rd} \\ \frac{1}{2}\left(\frac{(1-p)\bar{V}_{dd}}{1-\frac{p}{2}} - \bar{v}_{dd}\right), & \textit{otherwise} \end{cases} \tag{8}$$

and

$$\bar{V}_{dr} = \begin{cases} v - \frac{c}{p}, & \textit{if} \quad v < \bar{v}_{rd} \\ \frac{1}{2}\left(\frac{(1-p)\bar{V}_{dd}+v}{1-\frac{p}{2}} - \bar{v}_{dd}\right), & \textit{otherwise} \end{cases} \tag{9}$$

Proof. See the Appendix. □

From inspection of Eqs. (8) and (9), one can see that an increase in the cost of experimentation or an increase in the probability the breakthrough occurs all decrease both continuation payoffs. We complete the section with a remark that is key for the remainder of the paper.

Lemma 3. *(i)* $\bar{V}_{dd}$ *and* $\bar{V}_{rd}$ *are continuous and increasing in* v *with* $\bar{V}_{rd}$ *being strictly increasing for* $v \geq \bar{v}_{rd}$*, (ii)* $\bar{V}_{dr}$ *is continuous everywhere, strictly increasing for* $v < \bar{v}_{rd}$ *and* $v > \bar{v}_{rd}$ *but exhibits a (downward) jump discontinuity at* $v = \bar{v}_{rd}$*.*

Proof. See the Appendix. □

Lemma 3 states that although $\bar{V}_{dd}$ and $\bar{V}_{rd}$ are continuous and increasing in v, $\bar{V}_{dr}$ exhibits a downward jump discontinuity at $v = \bar{v}_{rd}$. Intuitively, when the value of the product exceeds the threshold above which the laggard enters the race, competition reduces the continuation payoff of the leader. As we show below, this result has significant implications on the equilibrium stopping times and welfare.

□ **Both Firms in the Research Stage**. The next and last step in the characterisation of the symmetric equilibrium is to study the subgame in which both firms are in the research stage. We have shown above that, once a firm succeeds in the research stage, there exists a unique symmetric equilibrium for all parameter values, which we have characterised in Lemma 1 and Lemma 2.

Suppose therefore that both firms are in the research stage. We first characterise the equilibrium stopping time of any of the firms experimenting. The expected (static) payoff of firm i by experimenting at intensity x^i in a period when the belief is α, provided that firm j does not experiment, is

$$\Big\{\alpha p \bar{V}_{dr} - c\Big\} x^i \tag{10}$$

which defines a unique cutoff in the posterior belief below which experimentation is not profitable.[11] This cutoff is given by

$$\bar{\alpha} = \min\left\{1, \frac{c}{p\bar{V}_{dr}}\right\} \tag{11}$$

We now study the behaviour of $\bar{\alpha}$ as a function of v and obtain the following result.

Proposition 1. *The equilibrium stopping time $\bar{\alpha}$ is continuous and strictly decreasing for $v < \bar{v}_{rd}$ and $v > \bar{v}_{rd}$; it exhibits an upward jump discontinuity at $v = \bar{v}_{rd}$.*

Proof. Note that for $v < \bar{v}_{rd}$,

$$\frac{c}{p\bar{V}_{dr}} = \frac{c}{pv - c}$$

which, given Assumption 1, is less than one. The discontinuity in $\bar{\alpha}$ is an immediate consequence of the discontinuity in $\bar{V}_{dr}$. □

Proposition 1 is one of the main results of the paper. It states that the equilibrium stopping time is everywhere continuous and strictly decreasing but exhibits an upwards jump discontinuity at $v = \bar{v}_{rd}$. This discontinuity arises from the discontinuity of $\bar{V}_{dr}$ as discussed in Lemma 3.

A natural question that arises concerns the characterisation of equilibrium intensity of experimentation for $\alpha > \bar{\alpha}$.[12] To study this, let us write the (static) expected payoff of firm i by experimenting at intensity x^i in a period when the belief is α, provided that firm j experiments at intensity x^j:

$$\alpha(px^i(1 - px^j)\bar{V}_{dr} + (1 - px^i)px^j\bar{V}_{rd} + p^2x^ix^j\bar{V}_{dd}) - cx^i \tag{12}$$

To explain Eq. (12), note that if the state is good, with probability $px^i(1 - px^j)$ firm i succeeds and firm j fails, in which case the corresponding continuation payoff is $\bar{V}_{dr}$; with probability $(1 - px^i)px^j$ firm i fails and firm j succeeds in which case the corresponding continuation payoff is $\bar{V}_{rd}$; with probability $p^2x^ix^j$, both firms succeed and the corresponding continuation payoff is $\bar{V}_{dd}$. The last term is the cost of experimentation.

[11] Note that, for a given α, this is the maximum "static" payoff a firm can attain, i.e., the payoff if the firm was a "monopolist", where "static" refers to the fact the the continuation payoff is not taken into account.

[12] For $\alpha = \bar{\alpha}$, the unique symmetric equilibrium intensity of experimentation is zero.

To characterise the symmetric equilibrium intensity of experimentation, let us re-write Eq. (12) as

$$\underbrace{\Big\{\alpha[pV_{dr} - p^2x^j(\bar{V}_{dr} + \bar{V}_{rd} - \bar{V}_{dd})] - c\Big\}x^i}_{\text{Expected Payof from } x^i, x^j \geq 0} + \underbrace{apx^j\bar{V}_{rd}}_{\text{Expected Payoff from } x^i=0} \tag{13}$$

If $\alpha(p\bar{V}_{dr} - p^2(\bar{V}_{dr} + \bar{V}_{rd} - \bar{V}_{dd})) \geq c$, Eq. (13) is increasing in x^i for every x^j (and strictly increasing if the inequality is strict); if $\alpha(p\bar{V}_{dr} - p^2(\bar{V}_{dr} + \bar{V}_{rd} - \bar{V}_{dd})) < c$, Eq. (13) is increasing in x^i for $x^j \leq (p\bar{V}_{dr} - \frac{c}{\alpha})/p^2(\bar{V}_{dr} + \bar{V}_{rd} - \bar{V}_{dd})$ (and strictly increasing if the inequality is strict) and strictly decreasing for $x^j > (p\bar{V}_{dr} - \frac{c}{\alpha})/p^2(\bar{V}_{dr} + \bar{V}_{rd} - \bar{V}_{dd})$. In the latter case, in the unique symmetric equilibrium, the two firms experiment at a strictly-less-than-maximum intensity; each firm experiments at an intensity that makes the rival indifferent between experimenting or not.

We summarise the unique symmetric equilibrium in the following proposition.

Proposition 2. *Suppose that both firms are in the research stage; then, in the unique symmetric equilibrium, for every $\alpha \geq \bar{\alpha}$:*

$$\bar{x}_{rr}(\alpha) = \min\left\{1, \frac{p\bar{V}_{dr} - \frac{c}{\alpha}}{p^2(\bar{V}_{dr} + \bar{V}_{rd} - \bar{V}_{dd})}\right\} \tag{14}$$

By inspection of Eq. (14), one can see that an implication of Proposition 2 is that experimentation is higher the higher is α and the lower is c. This implies that firms might start the race by experimenting at a maximum intensity but might reduce their intensities as they become more pessimistic regarding the viability of the project. Furthermore, the intensity of experimentation crucially depends on whether $v \lesseqgtr \bar{v}_{rd}$. In particular, we know from Lemma 2 that for $v < \bar{v}_{rd}$, $\bar{V}_{rd} = 0$ as the laggard finds it unprofitable to continue experimenting; for $v > \bar{v}_{rd}$, $\bar{V}_{rd} = 0$ and hence reduces equilibrium experimentation as each firm expects the other firm to suffer the cost of experimentation.[13] This is a consequence of the fundamental *free-riding effect* that has been highlighted in the strategic experimentation literature (e.g., Bolton and Harris (1999), Keller, Rady, and Cripps (2005)): the benefit of any of the firms from not experimenting might be positive since, conditional on the rival experimenting, there is a positive externality from a potential release of good news.

The equilibrium stopping time is depicted in Figure 2: the horizontal axis represents v; the vertical axis represents α. The stopping time is everywhere continuously decreasing but exhibits an upward discontinuity at $v = \bar{v}_{rd}$.

Another implication of Propositions 1 and 2 is that the intensity of experimentation is not always increasing in the value of the product. To see this, consider a (α, v) combination in the shaded region A in Figure 2. In this region firms will stop experimenting as we explained above. Suppose now that during a recession the value of the product falls. For the same posterior belief, such a fall will be accompanied by an increase in experimentation as firms find it profitable to enter the race knowing that the leader will also be the winner of the race.

[13]Note that $\bar{V}_{rd}$ appears only in the denominator, which means that a higher $\bar{V}_{rd}$ decreases $\bar{x}_{rr}(\alpha)$.

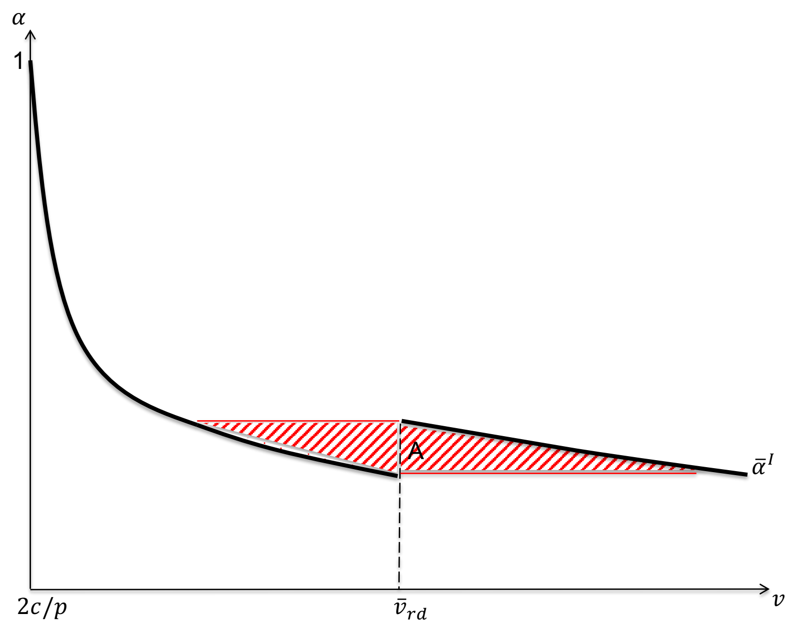


Figure 2

4 Welfare

We now describe the socially optimal policy and discuss in what ways the equilibrium differs from it and thus we examine the potential inefficiencies that might arise during the race.

First, note that, due to the assumption of no discounting, it is never socially optimal to invest in both technologies in the same period, in any of the stages. This is because the social cost of operating one technology, given that the other technology is active, is positive, whereas the social benefit is zero, i.e., the incremental value of succeeding in one technology is zero if the other technology succeeds. In fact, if both technologies are operated by a benevolent social planner, it is strictly dominant to experiment at a maximum intensity in at most one technology. Therefore, our objective is to examine the conditions under which it is socially optimal to operate one technology.

Suppose first that the research stage has been successfully completed. Whilst in the development stage, Assumption 1 implies that it always is socially optimal to experiment at a maximum intensity in exactly one technology.

The expected value of the product if one technology has advanced to the development stage is

$$V_d^* = v - \frac{c}{p} \tag{15}$$

Now suppose that no technology has advanced to the development stage. We wish to characterise the optimal stopping time, i.e., specify the unique threshold belief below which it is socially optimal to stop experimenting.

If the belief is α, the expected payoff of commiting to a plan of experimentation $\{x_\omega\}_{\omega=t}^T$

from period t till period T is given by the value function:

$$V_t^*(\alpha) = \alpha p x_t V_d^* - c x_t + (1 - \alpha p x_t) V_{t+1}^*(\alpha') \tag{16}$$

where α' is given in Eq. (1). It is only straightforward that if $\alpha p V_d^* - c > 0$, the planner will choose to experiment at a maximum intensity in period t, or $x_t = 1$. The optimal stopping time is characterised if in Eq. (16) one imposes that $V_t^*(\alpha) \geq 0$ and $V_{t+1}^*(\alpha') < 0$. In other words, the optimal stopping time is the last experiment such that the payoff is positive. The optimal stopping time is given by

$$\alpha^* = \min\left\{1, \frac{c}{p V_d^*}\right\} \tag{17}$$

In fact, one can characterise the unique threshold in the value of the product above which experimentation is socially optimal for a sufficiently high belief. This is specified in the following proposition.

Proposition 3. *The socially optimal stopping time α^* is everywhere continuously strictly decreasing. Moreover, $\alpha^* = \bar{\alpha}$ for $v < \bar{v}_{rd}$ and $\alpha^* > \bar{\alpha}$ for $v \geq \bar{v}_{rd}$.*

Proposition 3 states that the optimal stopping time is decreasing in v. Furthermore, it coincides with the equilibrium stopping time for $v < \bar{v}_{rd}$ but is strictly greater than the equilibrium stopping time for $v \geq \bar{v}_{rd}$. The latter is a result of the fundamental free-riding effect: even if it is socially optimal to continue experimenting firms stop prematurely.

Given the analysis thus far, it is not difficult to see that equilibrium experimentation is always inefficient. Nonetheless, different types of inefficiencies arise based on (i) the stage at which firms are, and, (ii) parameter values. To be more precise, we analyse the different types of inefficiencies by starting from the end and moving backwards. The following corollary summarises the possible inefficiencies when both firms have advanced to the development stage.

Corollary 1. *Suppose that both firms have qualified to the development stage; then both firms experiment at maximum intensity in equilibrium whilst it is socially optimal to experiment at a maximum intensity only in one technology.*

Corollary 1 is in line with the early contributions by Loury (1979), Lee and Wilde (1980) who show that firms may undertake excessive R&D because they do not internalise the negative externalities they impose on one another.

We can now study the subgame starting after one firm advances to the development stage while the rival firm is still in the research stage. We saw in Lemma 2 that there is a critical cutoff below which the firm in the research stage exits the race. The following corollarly highlights the possible inefficiencies.

Corollary 2. *Suppose that one firm has qualified to the development stage while the rival is still in the research stage; then, (i) for $v < \bar{v}_{rd}$ only the firm that is in the development stage experiments at a maximum intensity in equilibrium as is socially optimal, (ii) for $v \geq \bar{v}_{rd}$, it is socially optimal for only the firm that is in the development stage to expiriment at a maximum intensity but both firms experiment at a maximum intensity in equilibrium.*

Corollary 2 states that for products that are not sufficiently lucrative, advancement to the development stage by one firm is sufficient for the race to degenerate into a monopoly; for sufficiently lucrative products, duplication of investment arises as firms do not internalise the negative externalities they impose on one another. The former results is in line with early papers in the patent races literature (e.g., Fudenberg et al. (1983), Harris and Vickers (1985));[14] the latter result is similar to that obtained in Corollary 1.

Suppose now that both firms are in the research stage. As we argued in Proposition 3, $\bar{\alpha} \geq \alpha^*$ as $\bar{V}_{dr} \leq V_d^*$ with the inequality being strict for $v \geq \bar{v}_{rd}$.

Corollary 3. *Suppose that both firms are in the research stage; then, it is socially optimal to experiment in at most one technology at maximum intensity but in equilibrium both firms experiment.*

Corollary 3 states that apart from over-investment a potential under-investment might arise when both firms are in the research stage: for relatively high belief and/or value products, both firms experiment at a maximum intensity although it is socially optimal to experiment at a maximum intensity at most in one technology; for relatively low belief and/or value products, firms might experiment too little.

All the inefficiencies that characterised in Corollaries 1 and 2 are essentially *static* in the sense that they do not arise from the dynamics of the model. Unlike Corollaries 1 and 2, the inefficiency described in Corollary 3 is a *dynamic inefficiency*. The dynamic inefficiency is the result of the dynamics of the model: firms might under-invest due to the fundamental free-riding effect.

5 Discussion

■ **Model Relevance**. The model applies in innovation races in which completion of at least two stages is a prerequisite for a patent to be allocated and, at the outset, firms are uncertain about the overall feasibility of the project. For instance, if only one stage is completed, the requirements of the Patent Granting Authority are not satisfied. For instance in the pharmaceutical industry, the allocation of a patent is granted after a significant period has lapsed and several stages have been completed. One can assume that more than two stages (i.e., $N > 2$) are required before the product is successfully marketed but successful completion of the first two stages is sufficient for a firm to file an application for the allocation of a patent. In that case v denotes the expected monopoly profit minus the cost required to complete the $N - 2$ remaining stages.

□ **Asymmetric Equilibria**. Although we focus on the unique symmetric equilibrium, asymmetric equilibria exist. For instance, an asymmetric equilibrium entails only one firm experimenting at a maximum intensity for relatively low α instead of both firms experimenting. This equilibrium is more efficient than the symmetric equilibrium as it avoids the duplication of investment.

[14] This is sometimes called "ϵ-preemption".

□ **Robustness**. Although, we have imposed a number of simplifying assumptions, our results are robust in various alternative specifications. Below we discuss some:

- *Discounting*: Throughout, we have assumed that firms do not discount the future. This assumption was only for simplicity; all our results are valid even if firms discount the future. The most important implication of discounting regards the socially efficient stopping time. In particular, under discounting, the planner faces an inter-temporal trade-off: experimenting at maximum intensity in both technologies is costly but increases the probability of success. Hence, an impatient social planner would prefer to invest at a maximum intensity in both technologies. For the reasons explained above, under no discounting the planner faces no inter-temporal trade-off and hence would operate at most one technology at a time.

-*Continuous vs. discrete time*: Following Bolton and Harris (1999), most of the literature on strategic experimentation employs a continuous time framework.[15] Instead, we study a discrete-time counterpart of Keller, Rady, and Cripps (2005) exponential model, although our results are robust even in continous time.

-*Other issues*: We have assumed that the same bandit technology is used in both stages. This implies that (i) conditionally on the technology being armed in a good state, the probabilities of success in both stages are the same and (ii) the development stage is certain. Extension to an environment in which firms have different (symmetric) technologies in the two stages is straightforward and would not alter the results.

□ **Avenues for Future Research**. One interesting avenue for future research is to study a game in which firms do not perfectly observe the actions and outcomes of one another but they can voluntarily disclose the intermediate breakthrough as in Yildirim (2005), Gill (2008), Rieck (2010) and Akcigit and Liu (2015). In our model, when $v < \bar{v}_{rd}$, firms have strong incentive to disclose any intermediate breakthrough as soon as possible to discourage the rival from continuing; when $v \geq \bar{v}_{rd}$, firms would like to hide intermediate breakthroughs to protect themselves from competition. Nonetheless, in either case, the equilibrium intensity of experimentation depends on the beliefs firms entertain regarding the relative position of the rival which makes the analysis considerably more complicated.

Another interesting extension is to assume that firms do not disclose intermediate breakthroughs but receive a private information signal –perhaps due to espionage– regarding the state of the rival. This would also introduce different incentives for experimentation.

6 Conclusion

In this paper we studied a winner-takes-all R&D race as a two-armed bandit model. Two firms that competed on the development of a new product had to pass two stages (Research & Development) before acquiring a patent. Firms possessed identical bandit technologies able to provide them with the breakthrough that was required to complete a stage. Nonetheless, at the outset, firms were uncertain whether their technologies were

[15]For an exception see Heidhues, Rady, and Strack (2015).

armed in a good state, in which case the required breakthrough could occur with a strictly positive probability, or in a bad state, in which case the breakthrough could never occur. We assumed that firms observed one another's actions and outcomes.

We analysed the unique symmetric equilibrium of this dynamic game. We showed that for relatively low-value products firms experimented at a maximum intensity and the equilibrium stopping time coincided with the socially efficient stopping time; for relatively high-value products, free-riding could reduce experimentation and firms stopped experimenting prematurely. Perhaps surprisingly, we showed that a decrease in the value of the product could spur experimentation.

APPENDIX

■ **Proof of Lemma 1.** To find $\bar{V}_{dd}$, note that if both firms experiment at a maximum intensity every period, the continuation payoff of each firm is

$$\bar{V}_{dd} = \Big\{p(1-\frac{p}{2})v - c\Big\} + (1-p)^2\bar{V}_{dd}$$

Solving for $\bar{V}_{dd}$, one obtains that

$$\bar{V}_{dd} = \frac{p(1-\frac{p}{2})v - c}{1-(1-p)^2} = \frac{1}{2}\Big(v - \bar{v}_{dd}\Big)$$

where $\bar{v}_{dd} = 2c/p(2-p)$.

□ **Proof of Lemma 2.** To find $\bar{V}_{rd}$, note that, as we argued in the text, for $v < \bar{v}_{rd}$, the laggard exits the race; for $v \geq \bar{v}_{rd}$, the laggard continues in the race and experiments at a maximum intensity. The laggard's continuation payoff is given by

$$\bar{V}_{rd} = \Big\{p(1-p)\bar{V}_{dd} - c\Big\} + (1-p)^2\bar{V}_{rd}$$

which is equivalent to

$$\bar{V}_{rd} = \frac{p(1-p)\bar{V}_{dd} - c}{1-(1-p)^2} = \frac{p(1-p)\bar{V}_{dd} - c}{2p(1-\frac{p}{2})} = \frac{1}{2}\Big(\frac{(1-p)\bar{V}_{dd}}{1-\frac{p}{2}} - \bar{v}_{dd}\Big)$$

Regarding the leader's continuation payoff, for $v < \bar{v}_{rd}$, it is a monopolist and earns a continuation payoff

$$\bar{V}_{dr} = pv - c + (1-p)\bar{V}_{dr}$$

which is equivalent to $\bar{V}_{dr} = v - \frac{c}{p}$; for $v \geq \bar{v}_{rd}$, the laggard continues in the race and the leader's payoff is

$$\bar{V}_{rd} = \Big\{pv + p(1-p)\bar{V}_{dd} - c\Big\} + (1-p)^2\bar{V}_{dr}$$

which is equivalent to

$$\bar{V}_{rd} = \frac{pv + p(1-p)\bar{V}_{dd} - c}{1-(1-p)^2} = \frac{1}{2}\Big(\frac{(1-p)\bar{V}_{dd} + v}{1-\frac{p}{2}} - \bar{v}_{dd}\Big)$$

□ **Proof of Lemma 3.** To show that $\bar{V}_{dr}$ is discontinuous in v at $\bar{v}_{rd}$, it suffices to show that

$$v - \frac{c}{p} > \frac{1}{2}\Big(\frac{(1-p)\bar{V}_{dd} + v}{1-\frac{p}{2}} - \bar{v}_{dd}\Big) \Leftrightarrow$$

$$v - \frac{c}{p} > \frac{(1-p)\bar{V}_{dd} + v}{2-p} - \frac{1}{2}\bar{v}_{dd} \Leftrightarrow$$

$$\Big(1 - \frac{2}{2-p}\Big)v > \frac{(1-p)\bar{V}_{dd}}{2-p} + \frac{c}{p}\Big(1 - \frac{1}{2-p}\Big) \Leftrightarrow$$

$$v > \bar{V}_{dd} + \frac{c}{p} \Leftrightarrow$$

$$v - \frac{c}{p} > \frac{1}{2}v - \frac{c}{(2-p)p} \Leftrightarrow$$

$$v > \frac{c}{p} \cdot \frac{2(1-p)}{2-p} \Leftrightarrow$$

which is always satisfied.

References

Akcigit, Ufuk, and Qingmin Liu, "The role of information in innovation and competition," *Journal of the European Economic Association*, 14 (2015), 828–870.

Bednorz, J George, and K Alex Müller, "Possible high T_c superconductivity in the Ba- La-Cu- O system," *Zeitschrift für Physik B Condensed Matter*, 64 (1986), 189–193.

Bimpikis, Kostas, and Kimon Drakopoulos, "Disclosing information in strategic experimentation," *Manuscript, Grad. School of Bus., Stanford Uni.*, (2018).

Bimpikis, Kostas, Shayan Ehsani, and Mohamed Mostagir, "Designing Dynamic Contests," *Manuscript, Grad. School of Bus., Stanford Uni.*, (2015).

Bolton, P., and C. Harris, "Strategic experimentation," *Econometrica*, 67 (1999), 349–374.

Che, Yeon-Koo, and Johannes Horner, "Optimal design for social learning," *Manuscript, Dept. Econ., Columbia University*, (2015).

Choi, J., "Dynamic R&D competition under" hazard rate" uncertainty," *RAND Journal of Economics*, (1991), 596–610.

Dasgupta, P., and J. Stiglitz, "Uncertainty, industrial structure, and the speed of R&D," *Bell Journal of Economics*, (1980), 1–28.

DiMasi, J. A., "Success rates for new drugs entering clinical testing in the United States," *Clinical Pharmacology & Therapeutics*, 58 (1995), 1–14.

DiMasi, J. A., H. G. Grabowski, and R. W. Hansen, "Innovation in the pharmaceutical industry: new estimates of R&D costs," *Journal of health economics*, 47 (2016), 20–33.

DiMasi, J. A., R. W. Hansen, and H. G. Grabowski, "The price of innovation: new estimates of drug development costs," *Journal of health economics*, 22 (2003), 151–185.

Fudenberg, D., R. Gilbert, J. Stiglitz, and J. Tirole, "Preemption, leapfrogging and competition in patent races," *European Economic Review*, 22 (1983), 3–31.

Gill, David, "Strategic disclosure of intermediate research results," *Journal of Economics & Management Strategy*, 17 (2008), 733–758.

Grossman, G., and C. Shapiro, "Dynamic R & D competition," *The Economic Journal*, 97 (1987), 372–387.

Halac, Marina, Navin Kartik, and Qingmin Liu, "Contests for experimentation," *Journal of Political Economy*, 125 (2017), 1523–1569.

Harris, C., and J. Vickers, "Perfect Equilibrium in a Model of a Race," *The Review of Economic Studies*, 52 (1985), 193–209.

——, "Racing with uncertainty," *Review of Economic Studies*, 54 (1987), 1–21.

Heidhues, Paul, Sven Rady, and Philipp Strack, "Strategic experimentation with private payoffs," *Journal of Economic Theory*, 159 (2015), 531–551.

Keller, G., S. Rady, and M. Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, 73 (2005), 39–68.

Kremer, Ilan, Yishay Mansour, and Motty Perry, "Implementing the "wisdom of the crowd"," *Journal of Political Economy*, 122 (2014), 988–1012.

Lee, T., and L. Wilde, "Market structure and innovation: a reformulation," *Quarterly Journal of Economics*, 94 (1980), 429–436.

Loury, G., "Market structure and innovation," *Quarterly Journal of Economics*, (1979), 395–410.

Malueg, D., and S. Tsutsui, "Dynamic R&D competition with learning," *RAND Journal of Economics*, (1997), 751–772.

Mason, Robin, and Juuso Valimaki, "Learning in Contests," *Manuscript, Grad. School of Bus., Uni. of Exeter*, (2010).

Moscarini, Giuseppe, and Francesco Squintani, "Competitive experimentation with private information: The survivor's curse," *Journal of Economic Theory*, 145 (2010), 639–660.

Reinganum, J., "Dynamic Games of Innovation." *Journal of Economic Theory*, 25 (1981), 21–41.

——, "A dynamic game of R and D: Patent protection and competitive behavior," *Econometrica*, (1982), 671–688.

Rieck, Thomas, "Information disclosure in innovation contests," *Bonn Econ Discussion Papers*, (2010).

Vidali, Gianfranco, *Superconductivity: the next revolution?* (Cambridge university press, 1993).

Wu, Maw-Kuen, Jo R Ashburn, ClJ Torng, Ph H Hor, Rl L Meng, Lo Gao, Z Jo Huang, YQ Wang, and aCW Chu, "Superconductivity at 93 K in a new mixed-phase Y-Ba-Cu-O compound system at ambient pressure," *Physical review letters*, 58 (1987), 908.

Yildirim, Huseyin, "Contests with multiple rounds," *Games and Economic Behavior*, 51 (2005), 213–227.