Two Ways to the Top: Patent Races with Multiple Innovation Technologies

Michael Andrews*

June 10, 2016

Abstract

I build a duopoly model in which firms compete on a quality ladder. The only payoff-relevant state variable is the number of steps, or gap, between firms on the ladder. At each stage, each firm decides how much R&D effort to exert as well as which of two "innovation technologies" to use. A "standard innovation technology" has a high probability of success, but the cost for the follower increases as it falls more steps behind the leader. A "crazy innovation technology" has a lower probability of success, but the cost does not depend on the gap between firms. Thus, when the gap becomes sufficiently large, the follower switches to use the crazy innovation technology. If the follower has a single success using either innovation technology, it leapfrogs the leader. When only the standard innovation technology is available, the equilibrium exhibits familiar escape competition and discouragement effects: both firms exert the highest effort when the gap is small and lower effort when the gap is large. When both innovation technologies are available, changing costs in one state trickle down and affect firms' choices in other states in nontrivial ways. For instance, reducing the cost of the crazy innovation technology increases the follower's effort when the gap is large but decreases effort when the gap is small because it is less costly for followers to fall far behind and less profitable for leaders to move far ahead. Other comparative static results are illustrated with numerical simulations.

*Department of Economics, University of Iowa. *Email*: michael-j-andrews@uiowa.edu.

1 Introduction

Understanding how new ideas are developed is crucial for promoting innovation. While a large literature seeks to understand why firms pursue different types of innovations, the question of why firms utilize different strategies when racing to introduce the same innovation has largely been ignored.

Consider an environment in which firms or individuals race to introduce higher quality versions of a product, as if on a quality ladder. In many cases, there is a well-defined approach or set of techniques that firms and individuals follow to introduce the newest frontier-quality product. But sometimes innovators buck these well-worn conventions and try radically different, often "off-the-wall" strategies. Mokyr (1990, 252) colorfully refers to these instances as cases when "a crackpot hits the jackpot." The question is why innovators would even conceive of trying crackpot strategies in the first place when seemingly safer, traditional strategies can achieve the same outcome.

A concrete example may be useful. Ever since the discovery of superconductivity in 1911, scientists raced to find materials that superconducted at higher and higher temperatures.[1] In the field of superconductivity, "higher quality" was thus easy to measure: prestige went to the researchers who could find the highest transition temperature. In the mid 1980s, the consensus of the scientific community was that higher temperature superconductors would be found with organic compounds. Instead of following this conventional approach, two scientists working in an IBM lab in Zurich decided to experiment with perovskites.[2] Bednorz and Müeller's compound superconducted at 30K, a far higher temperature than the previous record. Blundell (2009, p. 107) explains what made this discovery so remarkable:

> There were several ways in which Bednorz and Müeller did not fit the typical pattern that might be expected for scientists making an astounding breakthrough. For example, astounding bursts of insight are often associated with a young genius whose mind is unencumbered with decades of familiarity with the accepted

[1] The background on superconductivity is drawn from Blundell (2009) and Matricon & Waysand (1994), both of which are very readable even to an economist.

[2] Perovskites are compounds of the form ABO_3, where O is oxygen and A and B are metal atoms. The Zurich IBM lab was known for its work on microscopes and was considered a backwater place to work on superconductivity.

> views...however, Alex Müeller was in his late fifties while doing his pioneering work on superconductivity... The surprise here is that they...were working very much against mainstream opinion...on a project which had no likelihood of success; it was later commented that if their original research proposal had been submitted to a university funding agency in the early 1980s, it would have been unlikely to receive a grant.

While the use of perovskites turned out to be an important breakthrough in superconductivity, at the time they were conducting their research it is not clear why Bednorz and Müeller would experiment with them when organic compounds were both more conventional and more promising. Bednorz and Müeller's seemingly crazy use of perovskites is more understandable when one realizes that neither was an expert in organic compounds. As the use of organics in superconductivity intensified, it became increasingly costly in terms of time and resources for the two researchers to master this separate branch of knowledge. So, as the researchers fell further behind the research frontier, attempting a crazy idea became relatively less costly. I build a simple theoretical model capturing this intuition. Even such a simple setup delivers non-trivial results: introducing the opportunity to work on crazy ideas when far below the frontier alters how much effort researchers exert when they are close to the frontier.

Throughout this paper, I refer to the techniques and strategies innovators use as "innovation technologies." More formally, an innovation technology is a menu of costs and arrival rates of success for a given innovation. I refer to the conventional strategies that the prevailing wisdom believes will be successful with high probability as a "standard innovation technology." I refer to the crackpot techniques as "crazy innovation technologies." Importantly, the costs of different innovation technologies may differ for innovators under different circumstances. The goal of this paper is not only to show under which circumstances the crazy innovation technology is optimal but also how firms strategically choose their investments to influence the circumstances under which their rivals operate.

In this paper, I study a repeated patent racing model in which firms choose both how much R&D effort to exert and which innovation technology to use.[3] More specifically, I build

[3]While I refer to them as "firms" throughout the paper, the agents in this game can equivalently be

a duopoly model in which firms compete on a quality ladder. The firm that is at a higher step of the ladder is the leader, and the firm at the lower step is the follower. At every step of the ladder, each firm can choose to utilize either the standard or the crazy innovation technology. A success by the follower using either technology will cause it to leapfrog the current leader. A success by the leader advances it one step on the ladder and increases the gap between firms. The state of the game is completely summarized by this gap. A larger gap delivers a larger flow payoff to the leader. Increased R&D effort increases the arrival rate of success. Because firms choose R&D effort in addition to the innovation technology, the action space has both continuous and discrete elements. As the follower falls further behind the leader, its cost per unit of R&D effort to use the standard technology increases, capturing the idea that followers must replicate the knowledge, technical skills, and complementary assets of leading firms in order to leapfrog them.[4] Because the crazy innovation technology uses skills that are so different than standard techniques, the follower does not need to replicate the leader's knowledge and hence the cost to use the crazy innovation technology does not vary with the gap between firms. Follower firms thus face a tradeoff: using the crazy innovation technology means a lower probability of success but also potentially large cost savings.

This model thus draws on and formalizes insights from the literature on complementary assets, as reflected by the different costs to utilize certain techniques, to explain why firms choose different innovation technologies. Teece (1986) first articulated the importance of complementary assets to innovation. The subsequent literature argues that firms' complementary asset positions can explain the timing of firm innovation decisions or the types of products that firms introduce.[5] The role of complementary assets in this model is a natural extension of this literature that shows how firms' asset positions can further affect what happens inside the black box of the R&D process.

It is worth noting which theories *cannot* explain the choice of innovation technologies as described above. Most notably, the replacement effect first articulated by Arrow (1962)

---

thought of as researchers racing to achieve a new breakthrough, a la Bednorz and Müeller.

[4]Hereafter, I refer to the R&D cost per unit of effort simply as the cost.

[5]Mitchell (1989) and Tripsas (1997) discuss how complementary assets affect leader firms' responses to the introduction of radical technologies by smaller rivals. Tripsas & Gavetti (2000) and Benner & Tripsas (2012) use case studies to argue that complementary assets affect the choices managers make, including the choice of what innovation technology to use. Wu et al. (2014) also argue that complementary assets can determine innovation decisions.

shows that established firms have less of an incentive to innovate because any improvements will at least partially replace existing sales. Replacement effects are operative in my model, as described below. But while the replacement effect influences how much effort firms exert, alone it cannot explain the decision to use the crazy innovation technology. Consider the superconductivity example discussed above. A higher-temperature superconducting compound replaces the previous state of the art, whether it comes from a standard organic compound or the crazy perovskites. A key assumption in my model is that final consumers see a breakthrough produced by a standard innovation technology as just as good as one produced by a crazy innovation technology. This assumption also distinguishes my model from a growing body of literature that seeks to understand why different types of firms introduce different kinds of innovations.[6]

My model is related in spirit to Segerstrom (1991) and the literature that builds on this model.[7] In these models, firms can choose to either advance to the quality frontier by imitating existing technologies or leapfrog the frontier by innovating. Imitation and innovation can be thought of as different innovation technologies as described here as they have different costs and benefits depending on the existing market structure. I expand on this literature by introducing strategic interactions: when making decisions, firms take into account that effort today can induce their rival to use a different innovation technology in future periods.[8]

I analyze the model in three steps: first when only the standard innovation technology is available, next when only the crazy innovation technology is available, and finally when both innovation technologies may be used. When only the standard innovation technology is available, the model exhibits the familiar escape competition and discouragement effects found in several other models with repeated patent races.[9] As the follower falls further

[6]See, for example, Acemoglu & Cao (2015), in which different types of innovations advance the frontier by different distances, and Akcigit et al. (2014) in which the payoff for innovation success depends on how many existing products lines the innovation can be applied to.

[7]See, for example, Davidson & Segerstrom (1998) and Cheng & Tao (1999), which explore the effect of non-linearities and policies that alter the costs and payoffs of imitation technologies.

[8]This strategic interaction also distinguishes my model from papers such as Rosen (1991) and Akcigit & Kerr (2015). In Rosen (1991), the model is static so actions in future states do not enter into firms' decisions. In Akcigit & Kerr (2015), the type of innovation technology used depends on how many product lines a firm operates; there is no strategic interaction between firms in the same product line.

[9]See, most famously, Aghion et al. (1997), Aghion et al. (2001), and Aghion et al. (2005), the latter of

behind, the cost of R&D increases at a sufficiently fast rate that it dominates any preemption incentives, and so the follower becomes discouraged and reduces its R&D spending. The leader, facing decreasing marginal returns from advancing an additional step and less risk of being leapfrogged, also reduces its effort as it gets further ahead. When the quality gap is small, R&D effort is the greatest as firms attempt to escape competition.[10] A model exhibiting both of these effects while still allowing leapfrogging is, as far as I know, novel to the literature.

How does the addition of the crazy innovation technology change incentives on the quality ladder? The crazy technology can be thought of as an "outside option": if the follower falls far behind, it always has the option to try a crazy idea. When the gap is sufficiently large, the follower begins using the crazy innovation technology; I refer to this as the switching state. Once the follower switches to using the crazy technology, it never wants to switch back until it has a success and leapfrogs its rival. After the follower switches to the crazy technology, its effort choice no longer depends on the size of the quality gap. This removes the discouragement effect when the follower is far behind. Because the follower's effort does not depend on the size of the gap, the leader can no longer discourage the follower through successful innovation; any change in the leader's effort after the follower switches is driven entirely by changes in flow payoffs.

Importantly, the option to utilize the crazy innovation technology affects firms' decisions before the switching state occurs as well. I show this first by comparing the case with both innovation technologies to a hypothetical case in which only the standard innovation technology is available. I then perform various comparative statics experiments in the case with both innovation technologies using numerical simulations. Not surprisingly, decreasing the cost of the crazy innovation technology increases the follower's effort and value after switching. By making these states more desirable, such a reduction in cost also reduces the discouragement effect: follower's do not need to exert as much effort to prevent falling

which argues that there is empirical evidence for these relationships.

[10] In an environment with step-by-step innovations, Hörner (2004) shows that the equilibrium need not exhibit greatest effort when the gap is small. When the lead is moderate, it is unclear whether high or low effort is desired, since the leading firm has neither an urgent need to defend its lead nor discourage the follower. Two assumptions of my model ensure greatest effort when the gap is small: leapfrogging and increasing costs. Increasing costs mean that the follower gets discouraged as it falls further behind, while leapfrogging guarantees that the leader always has an incentive to discourage its rival.

far behind the leader. Moreover, because a follower using the crazy innovation technology exerts more effort when costs are lower, inducing the follower to use that technology is less valuable to the leader, and therefore the leader exerts less effort to gain a big lead. This means that decreasing the costs of the crazy innovation technology, which only has a direct effect when the gap is large, reduces effort by both the leader and the follower when the gap is small.[11] An implication of this result is that, when the cost reduction is small, firms will spend less time using the crazy innovation technology when it is cheaper. This is because with a cheaper crazy technology, firms exert less effort when the gap is small and therefore spend more time in small-gap states. Moreover, after the follower switches to use the crazy technology, it exerts more effort and leaves that state more quickly. I also experiment with a reduction in the cost of the standard innovation technology. The effect of such a cost reduction on effort when the gap is large depends on whether it increases or decreases the follower's preemption incentives, which in turn depends on the particular parameter values of the model. These results suggest that the option to switch and use a crazy innovation technology affects repeated patent races in interesting and nontrivial ways.

This paper is organized as followers. The model is presented in Section 2. Section 2.1 presents the basic environment. I first examine the equilibrium when only the standard innovation technology is used in Section 2.2. In Section 2.3 I examine the equilibrium when only the crazy innovation is used. Finally, I examine the equilibrium when both are used in Section 2.4. I then use numerical simulations to illustrate some properties of this model in Section 3. Section 4 concludes.

2 Model

2.1 Environment

Two firms engage in competition along a quality ladder. Let one firm, call it firm A, be at level n_A on the quality ladder. Let its rival, firm B, be at level n_B. Define the quality gap $n \equiv n_A - n_B$. It turns out that in what follows n is the only payoff-relevant state variable.

[11] Such a phenomenon is an example of the "trickle down" effect identified by Acemoglu & Akcigit (2012).

Thus, without loss of generality, say the leader firm is at level n and the follower is at level $-n$. I drop the subscripts A and B.

Time is continuous. Firms are risk neutral and maximize profits. A flow profit term, Π_n, succinctly captures the reduced form of competition among firms as a function of the quality gap between them. When both firms are at the same step of the quality ladder, firms compete a la Bertrand and thus receive zero profits. When one firm is ahead of its rival, however, it produces a higher quality good and thus earns positive profits following standard limit pricing arguments. A higher quality product leads to strictly greater profits, but consumers' willingness to pay for higher quality is decreasing with the quality gap. Thus we can summarize each firm's flow profits by

$$\Pi_n = \left\{ \begin{array}{ll} 0 & \text{if } n \leq 0 \\ \pi_n & \text{if } n > 0 \end{array} \right\} \tag{1}$$

where π_n has the properties that $\forall\, n > 0$, $\pi_{n+1} > \pi_n$, $\pi_{n+1} - \pi_n < \pi_n - \pi_{n-1}$, and $\lim_{n\to\infty} \pi_n = \bar{\pi}$. Notice that every successful innovation by either firm is drastic in the sense of Reinganum (1983) and Reinganum (1985): a successful innovation makes the innovator a monopolist. If the current leader is the innovator, the new innovation completely replaces the innovator's profits from its previous highest quality product.

Firms advance along the quality ladder through successful innovation. Here, unlike in standard models of step-by-step innovation, the follower can leapfrog the leader with one single breakthrough.[12] Thus, one success by the follower advances the follower $n + 1$ steps along the quality ladder. However, the leader can only advance the frontier quality by one step at a time. Assume that the leader obtains a perpetual patent one every frontier breakthrough it achieves, so that the follower *must* attempt a frontier innovation.[13] This leapfrogging assumption also means that once firms leave the neck-and-neck state, they will never return to it.

At every instant, firms simultaneously decide both which innovation technology to use how much R&D effort to exert. An innovation technology is a menu of functions $\{h(\cdot), C(\cdot)\}$

[12]See Aghion et al. (2001) and Aghion et al. (2005), as well as more general racing models that have the step-by-step property, such as Budd et al. (1993) and Hörner (2004).

[13]This rules out what Hörner (2004) refers to as "frontier-hugging" strategies.

that respectively determine the arrival rate of success and cost per unit of effort, as described below. There are two possible innovation technologies. The first innovation technology is called a *standard* innovation technology, that is, an innovation technology that uses familiar techniques and approaches innovation in a predictable way. In what follows I denote this innovation technology with the superscript S. The second is called a *crazy* innovation, and so I denote it by the subscript C. These two innovation technologies vary in both their cost and arrival rate of success per unit of effort. Unlike in Rosen (1991) and Acemoglu & Cao (2015), both the standard and crazy innovations are drastic, meaning a single success with either technology advances the frontier by one step.

For both innovation technologies, success is stochastic. Higher R&D effort increases the Poisson arrival rate of success. With the standard innovation technology, for every R&D effort level x, the arrival rate is given by the function $h(x)$ with the following properties:

- $h'(x) > 0 \; \forall \; x$
- $h''(x) < 0 \; \forall \; x$
- $h(0) = 0$
- $h'(0) < 1$
- $h(x) < \bar{h} \; \forall \; x$

The first three properties are standard. The fourth property says that there is no Inada condition.[14] The fifth condition ensures that there is an upper bound on the arrival rate of success. Notice that the probability of both the leader and follower having success in time interval Δt is $o(\Delta t)$, and so is ignored throughout.

Because it does not take advantage of existing knowledge and expertise, the crazy innovation technology is successful with a lower arrival rate per unit of effort, $h^C(x) = \alpha h(x)$ for $\alpha \in [0, 1]$.

[14]In fact, the fourth condition is stronger than that. It, along with the strict concavity assumption, says that the arrival rate always increases slower than R&D effort. This assumption is important for the proof of Proposition 1.

With the standard innovation technology, each firm must pay flow cost $C_n^S(x)$ per unit of R&D effort. Notice that this flow cost depends on the quality gap. For simplicity, let the flow cost be $C_n^S * x$. The cost function is summarized by

$$C_n^S(x) = \left\{ \begin{array}{ll} \bar{c}x & \text{if } n \geq 0 \\ c_n^S x & \text{if } n < 0 \end{array} \right\} \tag{2}$$

where $\forall \ -n < 0$, $c_{-(n+1)}^S > c_{-n}^S$, $c_{-(n+1)}^S - c_{-n}^S > c_{-n}^S - c_{-(n-1)}^S$, and $c_{-1}^S > \bar{c}$. The intuition is that advancing the quality frontier by one step always has the same cost for any given level of R&D effort. The further behind a firm falls, however, the more costly it is to purchase each unit of R&D effort. This captures the fact that firms that are far behind must spend more to develop the knowledge, technical skills, or complementary assets necessary to introduce a frontier quality product. Note that since flow profits and costs depend only on the state variable n, each firm's problem is stationary for a given quality gap. An additional assumption on the "convexity" of c_{-n}^S is required:[15]

Assumption 1. $\frac{c_{-1}^S}{\bar{c}}(c_{-1}^S - \bar{c}) \geq \frac{\bar{\pi}h'(0)}{r}$.

Because the crazy innovation technology deviates so far from the existing knowledge base, a firm that tries to use it does not need to replicate existing knowledge or skills. Thus, the cost of the crazy innovation technology does not depend on the quality gap: $c_n^C = c^C \ \forall \ n \in (-\infty, \infty)$.

For simplicity, assume that firms can only work with one innovation technology at a time. Define the firm's value with quality gap n by

$$V_n(x_n, x_{-n}) = \max\{V_n^S(x_n, x_{-n}), V_n^C(x_n, x_{-n})\}. \tag{3}$$

In what follows, I drop the arguments (x_n, x_{-n}) to save on space. The continuous time

[15] This assumption guarantees that costs are increasing fast enough to overcome the follower's preemption incentive, as described below. However, this is a rather extreme sufficient condition. In the numerical analyses presented below, I relax this assumption and find no deviations from the equilibrium as described below.

Bellman equation for the leader that uses the standard innovation technology is given by[16]

$$rV_n^S = \max_{x_n}\{\pi_n - \bar{c}x_n + h(x_n)[V_{n+1} - V_n^S] + h(x_{-n})[V_{-1} - V_n^S]\}. \tag{4}$$

The Bellman equation for a leader that uses the crazy innovation technology is

$$rV_n^C = \max_{x_n}\{\pi_n - c^C x_n + \alpha h(x_n)[V_1 - V_n^C] + h(x_{-n})[V_{-1} - V_n^C]\}. \tag{5}$$

The Bellman equation for a follower or neck-and-neck firm that uses the standard innovation technology is

$$rV_{-n}^S = \max_{x_{-n}}\{-c_{-n}^S x_{-n} + h(x_{-n})[V_1^S - V_{-n}^S] + h(x_n)[V_{-(n+1)} - V_{-n}^S]\} \tag{6}$$

The continuous time Bellman equation for a follower who uses the crazy innovation technology is given by

$$rV_{-n}^C = \max_{x_{-n}}\{-c^C x_{-n} + \alpha h(x_{-n})[V_1^S - V_{-n}^C] + h(x_n)[V_{-(n+1)} - V_{-n}^C]\}. \tag{7}$$

I examine three cases below. First, the parameters are such that firms never want to use the crazy innovation technology. Second, the parameters are such that the follower always uses the crazy innovation technology. Finally, I consider the case when both the standard and crazy innovation technologies will be used, depending on the state.

2.2 Standard Innovation Technology

Assume first that $\alpha = 0$ so that the crazy innovation technology never yields success with positive probability. Hence firms will never choose the crazy innovation technology when they exert positive effort, and so here I consider the standard innovation technology in isolation. So the leader's and follower's continuous time Bellman equation is always given by (4) and (6), respectively.

Taking first order necessary conditions of these Bellman equations with respect to the

[16]See Appendix B.1 for a derivation of the Bellman equations.

firm's own effort is straightforward and gives the following optimal R&D effort levels:

$$x_n^* = \max\{h^{'-1}(\frac{\bar{c}}{V_{n+1}^S - V_n^S}), 0\}, \tag{8}$$

$$x_0^* = \max\{h^{'-1}(\frac{\bar{c}}{V_1^S - V_0^S}), 0\}, \tag{9}$$

and

$$x_{-n}^* = \max\{h^{'-1}(\frac{c_{-n}}{V_1^S - V_{-n}^S}), 0\}. \tag{10}$$

These, along with Assumption 1, is sufficient to characterize the equilibrium of this game.

Proposition 1. *A Markov perfect equilibrium exists and has the following properties:*

a) $0 \leq V_n^S \leq \frac{\bar{\pi}}{r} \; \forall \; n$

b) $V_1^S > V_0^S$

c) $V_m^S > V_{-n}^S \; \forall \; m, n > 0$

d) $x_{-(n+1)}^* \leq x_{-n}^* \; \forall - n < 0$

e) $\exists \; -\check{n}$ *such that* $x_{-n}^* = 0 \; \forall \; -n \leq -\check{n}$

f) $V_n^S > V_{n-1}^S \; \forall \; n \geq 0$

g) $V_{n+1}^S - V_n^S < V_n^S - V_{n-1}^S$

h) $\exists \; \hat{n}$ *such that* $x_n^* < x_{n+1}^* \; \forall n < \hat{n}$ *and* $x_n^* = x_{n+1}^* = 0 \; \forall \; n \geq \hat{n}$

i) $V_{-(n+1)}^S \leq V_{-n}^S \; \forall \; -n \leq 0$

Proof. See Appendix B.2 □

Assumption 1 is sufficient to guarantee that the follower's effort is monotonically decreasing as n increases. The follower faces two countervailing incentives, as seen in equation (10). First, as the follower falls further behind it faces a higher flow cost per unit of effort, which leads it to reduce its effort as n increases. At the same time, falling further behind makes becoming the new leader relatively more attractive, which increases the follower's preemption incentive and induces it to increase its effort. Assumption 1 ensures that the discouragement effect dominates the preemption effect. Once monotonicity of the follower's effort is established, it is relatively straightforward to show the other properties.

The result is an equilibrium that delivers the familiar escape competition and discouragement effects: R&D effort is most intense when firms are neck-and-neck; as the leader gets farther ahead, the marginal benefit from advancing an additional step falls, and so effort decreases; and the follower becomes discouraged as the gap increases and so decreases its effort.[17] In the step-by-step innovation literature a la Aghion et al. (2001) and Aghion et al. (2005), the quality gap-R&D effort relationship comes from the fact that, to pass the industry leader, followers must introduce several consecutive successful innovations. This is difficult, and so the follower becomes discouraged as it falls further behind. When the gap is small, firms work harder to escape competition and discourage their rival. In my model, increasing costs play the same role as the step-by-step assumption, discouraging the follower as the gap size increases. The leapfrogging assumption ensures that the leader always has an incentive to discourage its rival, although this effort decreases as the follower falls further behind and the marginal benefits of discouragement decrease.

If $\hat{n} \geq \check{n}$, then $\hat{n}$ is an absorbing state. That is, once the leader is $\hat{n}$ steps ahead of the follower, then neither firm exerts any effort and therefore the state never changes. If $\check{n} > \hat{n}$, then an absorbing state does not exist. For any gap size that is reached with positive probability, the follower always exerts positive effort, and the follower will eventually have some success, reverting the state to $n = 1$.

2.3 Crazy Innovation Technology

Now assume that $\alpha = 1$ and $\bar{c} < c^C < c^S_{-1}$. Now clearly the crazy innovation technology dominates whenever a firm is one or more steps behind, and the standard innovation tech-

[17] In this particular model, leapfrogging means that firms never return to the neck-and-neck state once they exit it. Nevertheless, I include the neck-and-neck state in the analysis because it is a natural initial condition of the game. Even ignoring the neck-and-neck state, the same intuition applies: R&D effort is higher when the gap is n steps than when it is $n + 1$ steps for any positive n.

nology is preferred otherwise. Thus we can assume that the follower always uses the crazy innovation technology.

Because of the leapfrogging assumption, a single success always advances the follower to the same state, V_1^S. Also recall that the cost per unit of effort is the same in every state when the follower uses the crazy innovation technology. Thus, it is intuitively clear that the follower's value function and choice of R&D effort does not depend on the state n when the follower uses the crazy innovation technology. This is shown formally in Proposition 2.

Proposition 2. *For all $-n < 0$, $x^*_{-n} = x^C$ and $V^C_{-n} = V^C$, where x^C and V^C are constants. Furthermore, $V_1^S > V_0^S > V^C$. In addition, $V^S_{n+1} > V^S_n$ and $x^*_{n+1} \leq x^*_n \; \forall \; n$, and $\exists \; \hat{n}$ such that $x^*_n = 0 \; \forall \; n \geq \hat{n}$. For $n < \hat{n}$, $x^*_{n+1} < x^*_n$.*

Proof. See Appendix B.3. □

The crazy innovation technology delivers an equilibrium identical to the fast-catchup environment in Acemoglu & Akcigit (2012) and Acemoglu (2009). The follower's effort is constant, while the leader's effort declines as its lead increases. Clearly, in such an environment there is no discouragement effect as the follower falls further behind because the marginal benefit of a successful innovation does not decrease as the gap gets larger. I explore the implications of this fact in more detail in section 2.4.

2.4 Multiple Innovation Technologies

Now assume that:

Assumption 2. *$c^C \geq c^S_{-1}$ and $\alpha \in (0, 1)$.*

Assumption 2 ensures that the follower prefers to work on the standard innovation technology in at least some states while the leader and firms in the neck-and-neck state always strictly prefer to work on the standard innovation. In Lemma 1, I show that the intuition established in Propositions 1 and 2 carries over to cases when the follower may choose either innovation technology. Namely, the follower's R&D effort and value while using the standard innovation technology decreases with the size of the quality gap while effort and value with the crazy innovation technology are constant. What this means is that once the follower switches to use the crazy innovation technology it never wants to switch back.

Lemma 1. *If* $\exists\ \tilde{n}$ *such that* $V^C_{-\tilde{n}} \geq V^S_{-\tilde{n}}$*, then* $V^C_{-n} \geq V^S_{-n}\ \forall\ -n < -\tilde{n}$.

Proof. See Appendix B.4 $\square$

Lemma 1 establishes that there is a cutoff gap-size $\tilde{n}$ such that the follower uses the standard innovation technology when the gap is smaller than $\tilde{n}$ and the crazy innovation technology when the gap is $\tilde{n}$ or larger. Then for $n \geq \tilde{n}$ the follower's Bellman equation is

$$rV_{-n} = rV^C = \max_{x^C}\{-c^C x^C + \alpha h(x^C)[V^S_1 - V^C]\}. \tag{11}$$

For $n = \tilde{n} - 1$, the follower's Bellman equation is

$$\begin{aligned} rV_{-n} = rV^S_{-(\tilde{n}-1)} = \max_{x_{-(\tilde{n}-1)}} \{ & -c^S_{-(\tilde{n}-1)} x_{-(\tilde{n}-1)} + h(x_{-(\tilde{n}-1)})[V^S_1 - V^S_{-(\tilde{n}-1)}] \\ & + h(x_{\tilde{n}-1})[V^C - V^S_{-(\tilde{n}-1)}]\}. \end{aligned} \tag{12}$$

Finally, for $n < \tilde{n} - 1$, the follower's Bellman equation is

$$rV_{-n} = rV^S_{-n} = \max_{x_{-n}}\{-c^S_{-n} x_{-n} + h(x_{-n})[V^S_1 - V^S_{-n}] + h(x_n)[V^S_{-(n+1)} - V^S_{-n}]\}. \tag{13}$$

When the follower uses the crazy innovation technology (that is, when $n \geq \tilde{n}$, its effort is given by

$$x_{-n} = x^C = \max\{h'^{-1}(\frac{c^C}{\alpha[V^S_1 - V^C]}), 0\}. \tag{14}$$

When the follower uses the standard innovation technology (when $n < \tilde{n}$), the follower's effort is given by

$$x_{-n} = \max\{h'^{-1}(\frac{c^S_{-n}}{V^S_1 - V^S_{-n}}), 0\}. \tag{15}$$

Finally, the leader's effort is written identically to Equation 8,

$$x_n = \max\{h'^{-1}(\frac{\bar{c}}{V^S_{n+1} - V^S_n}), 0\}. \tag{16}$$

The equilibrium thus qualitatively resembles the case with only the standard innovation technology when $n < \tilde{n} - 1$. For $n \geq \tilde{n}$, the follower's R&D effort choice is constant. For $n \geq \hat{n}$, the leader's R&D effort choice declines as described in Proposition 2. Once the follower has switched to the crazy innovation technology, the leader can no longer discourage the follower by increasing the size of the gap, and so the leader's only reason to exert R&D effort when $n \geq \tilde{n}$ is to achieve higher flow profits.

Clearly, the switch to the crazy innovation technology increases the arrival rate for the follower when the gap is very large.[18] While superficially similar, this result is fundamentally different than the "go-for-broke" effect in Anderson & Cabral (2007). In that case, followers work harder because they have nothing to lose; the game will end if the follower does not have an immediate success. Here, however, the follower switches so that it no longer has to worry about falling further behind; switching means that the game never ends. In this sense, the crazy innovation technology provides insurance so that the follower never has to worry about "go-for-broke" or end-of-game effects. But switching has one other effect: it removes the term $h(z^*_n)[V_{-(n+1)} - V_{-n}$ from the follower's Bellman equation. This removes part of the follower's preemption incentive: instead of working harder to prevent the end of the game, the follower's effort is motivated entirely by the rewards of leapfrogging. This fact has an immediate implication for the follower's effort choice after switching. The preemption incentive induces the follower to work harder; by removing this incentive, switching to the crazy innovation may actually entail a lower level of effort than in the case with just the standard innovation. Conditions in which this is the case are shown in Lemma 2.

Lemma 2. *i) Define x^C as the follower's equilibrium choice of effort on the crazy innovation in the switching state. Define $x^*_{-\tilde{n}}$ as the follower's equilibrium choice of effort on the standard innovation in the switching state. There are three cases:*

[18] Recall that with the standard innovation technology, the follower's effort goes to zero as the gap gets large, while with the crazy innovation technology effort is constant.

1. *If* $\frac{c^C}{c^S_{-\tilde{n}}} \leq \frac{\alpha r[c^C - \bar{c}]}{\bar{\pi} h'(0)}$*, then the* $x^C \geq x^*_{-\tilde{n}}$*.*

2. *If* $\frac{c^C}{c^S_{-\tilde{n}}} \geq \frac{\alpha \bar{\pi} h'(0)}{r[c^S_{-1} - \bar{c}]}$*, then* $x^C \leq x^*_{-\tilde{n}}$*.*

3. *If* $\frac{c^C}{c^S_{-\tilde{n}}} \in \left(\frac{\alpha r[c^C - \bar{c}]}{\bar{\pi} h'(0)}, \frac{\alpha \bar{\pi} h'(0)}{r[c^S_{-1} - \bar{c}]}\right)$ *then it is ambiguous whether* x^C *or* $x^*_{-\tilde{n}}$ *is larger.*

Proof. See Appendix B.5. □

3 Simulation Analysis

In this section, I investigate the comparative statics of this model using numerical simulations. In the simulations below, I ignore Assumption 1. While this assumption is useful for the proof of Proposition 1, it establishes a stringent sufficient condition. The examples below show that it is not necessary in order to have an equilibrium with the properties described above. ***I also relax the assumption that there exists an*** $\bar{x}$ ***such that*** $h'(x) = 0$ ***for any*** $x \geq \bar{x}$***. This assumption is useful to prove that there exists a gap sufficiently large that the leader never exerts any R&D effort, even without establishing any other properties of the equilibrium, however this assumption is also not necessary.***[19]

3.1 Comparing Environments with One and Two Innovation Technologies

I first compare the equilibrium established above to a hypothetical case in which only the standard innovation technology is possible. Such a comparison is instructive because most of the existing literature considers environments in which only one innovation technology can be used.

Some of the differences between the cases with one and two innovation technologies are obvious. For instance, when there is a state at which the follower strictly prefers to use the crazy innovation technology, then the follower must have a higher value in those states

[19] For computational simplicity, in the simulations below I assume that there is a maximum gap size. More specifically, I assume that the leader cannot be more than ten steps ahead of its rival. In most cases, $x^{S*}_n = 0$ for an n smaller than 10, so this assumption is without loss of generality. Even in the few cases when this is not the case, the dynamics I am interested in describing happen at smaller levels of n.

when both technologies are allowed. Since the follower's effort goes to zero as the gap gets large when using only the standard innovation technology, there must eventually be states in which the follower exerts greater effort and achieves success at a higher rate by using the crazy innovation technology; this lowers the leader's value in those states.

But some comparisons are not so straightforward. For example, how does the option to utilize the crazy innovation technology affect firms' effort choices and arrival rates around the switching state? In Figure 1, the follower switches to the crazy innovation technology when $n = -2$. At $n = -2$, the follower exerts more effort in the case with two innovation technologies, as described in Lemma 2. But each unit of effort is less effective when the follower switches, so the follower's equilibrium arrival rate is actually lower at $n = -2$ in the case with two innovation technologies; the crazy innovation technology is still preferred by the follower because it is less costly to use. The leader's effort and arrival rate is lower in the case with 2 innovation technologies for two reasons. First, notice that the follower's arrival rate is higher when it uses the crazy innovation technology for all $n = 3$ and larger. This makes advancing to those states less attractive to the leader. In other words, the crazy innovation introduces a negative "trickle-down" effect: because V_3 decreases, this lowers V_2 and hence x_2^{S*}, which in turn lowers V_1 and x_1^{S*}.[20] Second, in any given state, firms' arrival rates are strategic complements within that state. This means that the follower's lower arrival at $n = -2$ lowers the leader's optimal effort choice at $n = 2$.

The presence of trickle-down effects when two innovation technologies are available means that the relationship between the quality gap and innovative performance may be much more subtle than the simple relationship found when there is only one innovation technology. Fig. 1 shows how the effect of switching to the crazy innovation when the gap is large can attenuate arrival rates even when the gap is small. Such results suggest that understanding policies that change the relative costs and benefits for different types of firms or different types of innovative projects may not be straightforward. I examine some of these issues in the following subsections.

[20] Acemoglu & Akcigit (2012) use the term "trickle down" effect to describe how increasing the strength of a leader's intellectual property protection when the gap is large gives firm's an additional incentive to reach the large-gap states, and hence also affects incentives when the gap is small. While the context is different, the same incentives are at work here.

3.2 Changing the Costs of a Crazy Innovation

Here I consider a small decrease in the cost of using the crazy innovation technology. Fig. 2 illustrates the effects of such a small decrease. Intuitively, the lower cost of the crazy innovation technology increases the follower's effort and arrival rate after switching. Just as in Figure 1, the leader now has less of an incentive to advance $\tilde{n}$ steps ahead, reducing the leader's effort.

An immediate implication of this finding is that, when the costs of using the crazy innovation technology are lower, firms will spend less total time utilizing the crazy technology. This is because with lower costs, the follower exerts more effort after switching and therefore more quickly advances the state to one where the gap is small. At the same time, the escape competition incentives are reduced, so firms advance to states in which the follower uses the crazy technology less often.

It is important to note that this intuition only describes what happens when the decrease in costs is sufficiently small that it does not change the state at which the follower switches to the crazy innovation. When the switching state does not change, it is easy to see that the follower's arrival rate is higher in the lower cost case than the higher cost case in the switching state. If the switching state changes, however, say from $-n*$ to $-(n*-1)$, then the follower's effort in the low cost case need not be higher then the high cost case in state $-(n*-1)$.

3.3 Changing the Costs of a Standard Innovation

In this section, I alter the cost of the standard innovation technology. First, I lower the cost for the standard innovation only when $n \geq 0$. Recall that this means decreasing $\bar{c}$, while leaving c_n^S the same for all $n \leq 0$. Such a cost reduction is essentially a subsidy for the leader firm, since the leader always uses the standard innovation technology. This case is shown in Figure 3. Even though this change lowers the costs of the standard innovation technology only, it increases the follower's arrival rate while using the crazy innovation technology.[21] The reason for this is clear: by making it more attractive to become the leader, the follower

[21]The increase is very small, but it is present; the figures are drawn so that a change in effort of zero results in no visible line at all.

has a greater incentive to exert R&D effort, regardless of which innovation technology it uses. The lower cost for the leader also induces the leader to exert much more effort. If the follower were forced to use the standard innovation technology only, the increase in the leader's arrival rate in states when the gap is bigger would induce the follower to exert even more effort, since falling to those states is now less attractive. Because the follower's effort level does not depend on the gap size when it uses the crazy innovation technology, this effect is absent here.

Next, I consider a case in which the costs of the standard innovation technology are lowered proportionally for every n. This case is shown in Figure 4. Whereas in the previous paragraph lowering $\bar{c}$ led to more effort when the follower used the crazy innovation technology, here the opposite is the case: when c_n^S is lower for all n, the follower's arrival rate after switching to the crazy innovation is lower. Why might this be the case? By lowering $\bar{c}$, the direct effect of lowering the cost is to make V_1^S more attractive, which should induce the follower to exert more effort at every n. But lowering all costs also lowers c_{-1}^S, so falling behind is no longer as costly. Moreover, if the current follower does retake the lead, it expects to hold it for a shorter period of time because $h(x_{-1}^{S*})$ has increased. Thus, the marginal benefit of becoming the leader is reduced. In the figure, this effect dominates, leading the follower to exert less effort when it is using the crazy innovation technology. Note that this need not be the case, however. Which effect dominates depends on the relative sizes of the cost decrease at different levels of n. In fact, it may be useful to think of the case in Figure 3 as the extreme case when all of the cost reduction goes to $\bar{c}$, while the case in Figure 4 presents another extreme when the cost reduction is distributed proportionally to all levels of n.

In total, the results from these simulations show that even simple comparative statics exercises are not straightforward in such a model. While the direct effect of a decrease in costs in a particular state n are easy to see, these costs have trickle down effects, changing relative payoffs in other states. These trickle down effects influence firms' choices of R&D effort and innovation technology in non-trivial ways.

4 Conclusion

In this paper, I build a duopoly model of a repeated patent race on a quality ladder where the payoff-relevant state variable is the gap between firms on the ladder. At each step, each firm chooses how much R&D effort to exert. For the follower, the cost per unit of effort to leapfrog the leader using the standard innovation technology increases with the size of the gap. In a baseline model with only the standard innovation technology, R&D effort is highest when the gap between firms is small. When the follower falls further behind, it becomes discouraged and reduces its effort.

When firms must choose not only R&D effort but also which innovation technology to use, the model becomes considerably more complicated. When the gap between firms is sufficiently large, the follower switches to using the crazy innovation technology. When the follower switches, its effort can be either higher or lower than its effort in the same state when only the standard technology is available. With both technologies, firms have less of an incentive to invest when the gap is small because falling far behind is no longer as costly. At the same time, by raising the follower's value in the worst possible states, the crazy innovation technology reduces the discouragement effect and hence increases the follower's effort in those states. These effects suggest that the relationship between competition and R&D effort is more subtle than simple one-innovation technology case. Understanding the effects of policies such as targeted R&D subsidies on innovation outcomes is not straightforward in this framework. The takeaway from this exercise is clear: the menu of innovation technologies available to firms matters for the outcome of patent races.

References

Acemoglu, D. (2009). *Introduction to modern economic growth.* Princeton: Princeton University Press.

Acemoglu, D., & Akcigit, U. (2012). Intellectual property rights policy, competition and innovation. *Journal of the European Economic Association*.

Acemoglu, D., & Cao, D. V. (2015). Innovation by entrants and incumbents. *Journal of Economic Theory*.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: an inverted-U relationship. *Quarterly Journal of Economics*.

Aghion, P., Harris, C., Howitt, P., & Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies*.

Aghion, P., Harris, C., & Vickers, J. (1997). Competition and growth with step-by-step innovation: an example. *European Economics Review*.

Akcigit, U., Hanley, D., & Serrano-Velarde, N. (2014). *Back to basics: basic research spillovers, innovation policy and growth.* (NBER Working Paper)

Akcigit, U., & Kerr, W. R. (2015). *Growth through heterogeneous innovations.* (NBER Working Paper)

Anderson, A., & Cabral, L. M. B. (2007). Go for broke or play it safe? Dynamic competition with choice of variance. *RAND Journal of Economics*.

Arrow, K. J. (1962). Economic welfare and the allocation of resources for invention. In S. S. R. Council (Ed.), *The rate and direction of inventive activity: Economic and social factors.* Princeton: Princeton University Press.

Benner, M. J., & Tripsas, M. (2012). The influence of prior industry affiliation on framing in nascent industries: the evolution of digital cameras. *Strategic Management Journal*.

Blundell, S. (2009). *Superconductivity: a very short introduction*. Oxford: Oxford University Press.

Budd, C., Harris, C., & Vickers, J. (1993). A model of the evolution of duopoly: does the asymmetry between firms tend to increase or decrease? *Review of Economic Studies*.

Cheng, L. K., & Tao, Z. (1999). The impact of public policies on innovation and imitation: the role of r&D technology in growth models. *International Economic Review*.

Davidson, C., & Segerstrom, P. (1998). R&D subsidies and economic growth. *RAND Journal of Economics*.

Federgruen, A. (1978). On N-person stochastic games with denumerable state space. *Advances in Applied Probability*.

Hörner, J. (2004). A perpetual race to stay ahead. *Review of Economic Studies*.

Matricon, J., & Waysand, G. (1994). *The cold wars: a history of superconductivity*. New Brunswick, NJ: Rutgers University Press.

Mitchell, W. (1989). Whether and when? Probability and timing of incumbents' entry into emerging industrial subfields. *Administrative Science Quarterly*.

Mokyr, J. (1990). *The lever of riches: technological creativity and economic progress*. Cambridge: Oxford University Press.

Reinganum, J. F. (1983). Uncertain innovation and the persistence of monopoly. *American Economic Review*.

Reinganum, J. F. (1985). Innovation and industry evolution. *Quarterly Journal of Economics*.

Rosen, R. J. (1991). Research and development with asymmetric firm sizes. *RAND Journal of Economics*.

Segerstrom, P. S. (1991). Innovation, imitation, and economic growth. *Journal of Political Economy*.

Teece, D. J. (1986). Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy*.

Tripsas, M. (1997). Unraveling the process of creative destruction: complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal*.

Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: evidence from digital imaging. *Strategic Management Journal*.

Wu, B., Wan, Z., & Levinthal, D. A. (2014). Complementary assets as pipes and prisms: innovation incentives and trajectory choices. *Strategic Management Journal*.

A Graphs

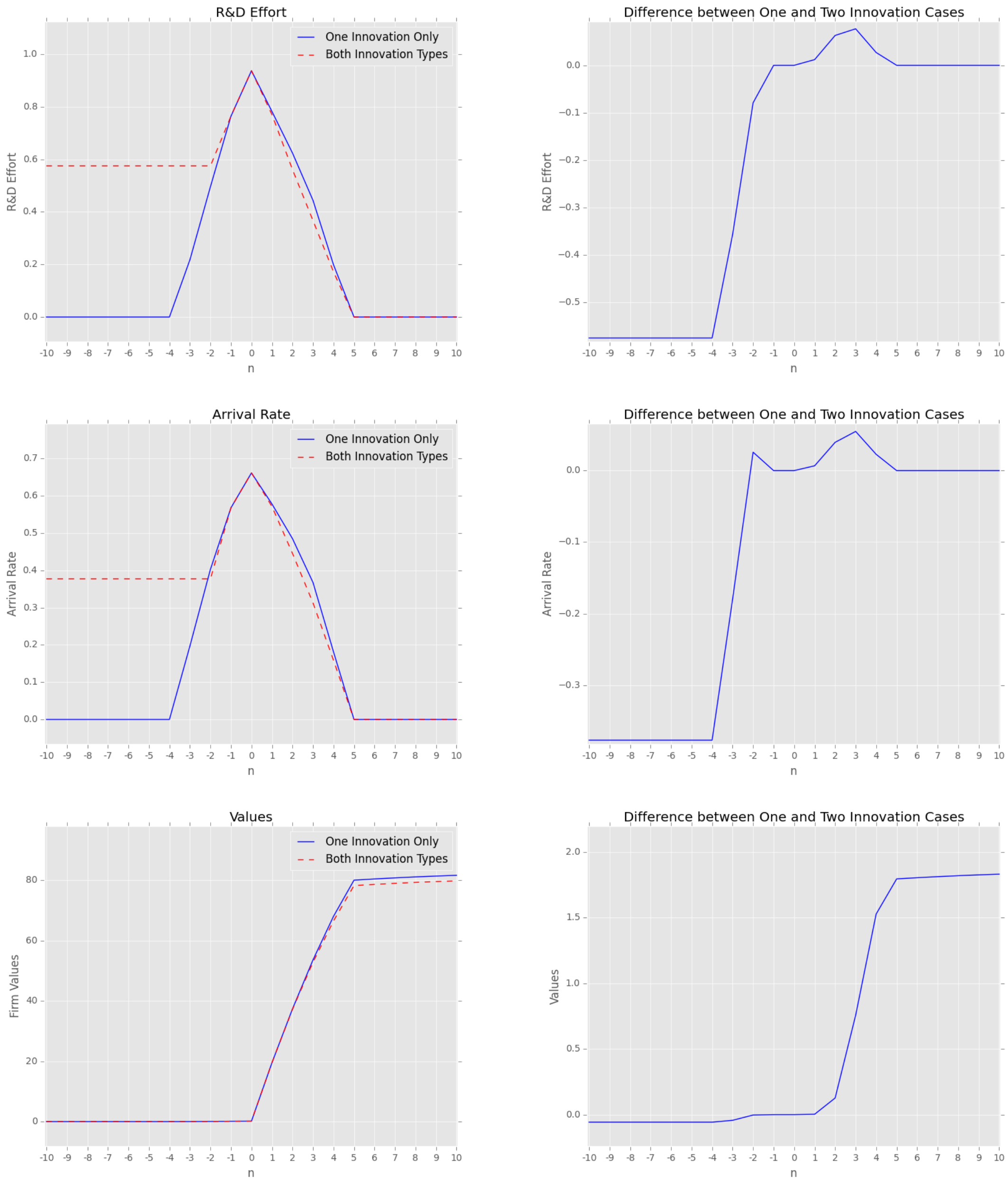

Figure 1: Compare the benchmark case when only the standard innovation technology is available to the case when both innovation technologies are available.
Parameter Values:
$h(x) = \alpha \ln(x+1)$, $\beta = .96$, $\pi_n = \{500, 950, 1350, 1700, 2000, 2010, 2019, 2027, 2034, 2040\}_{n=1}^{10}$, $c_{-n} = \{59, 50, 42, 35, 29, 24, 20, 16, 13, 11\}_{-n=-10}^{-1}$, $\bar{c} = 10$, $c^R = 10.25$, $\alpha = .83$.

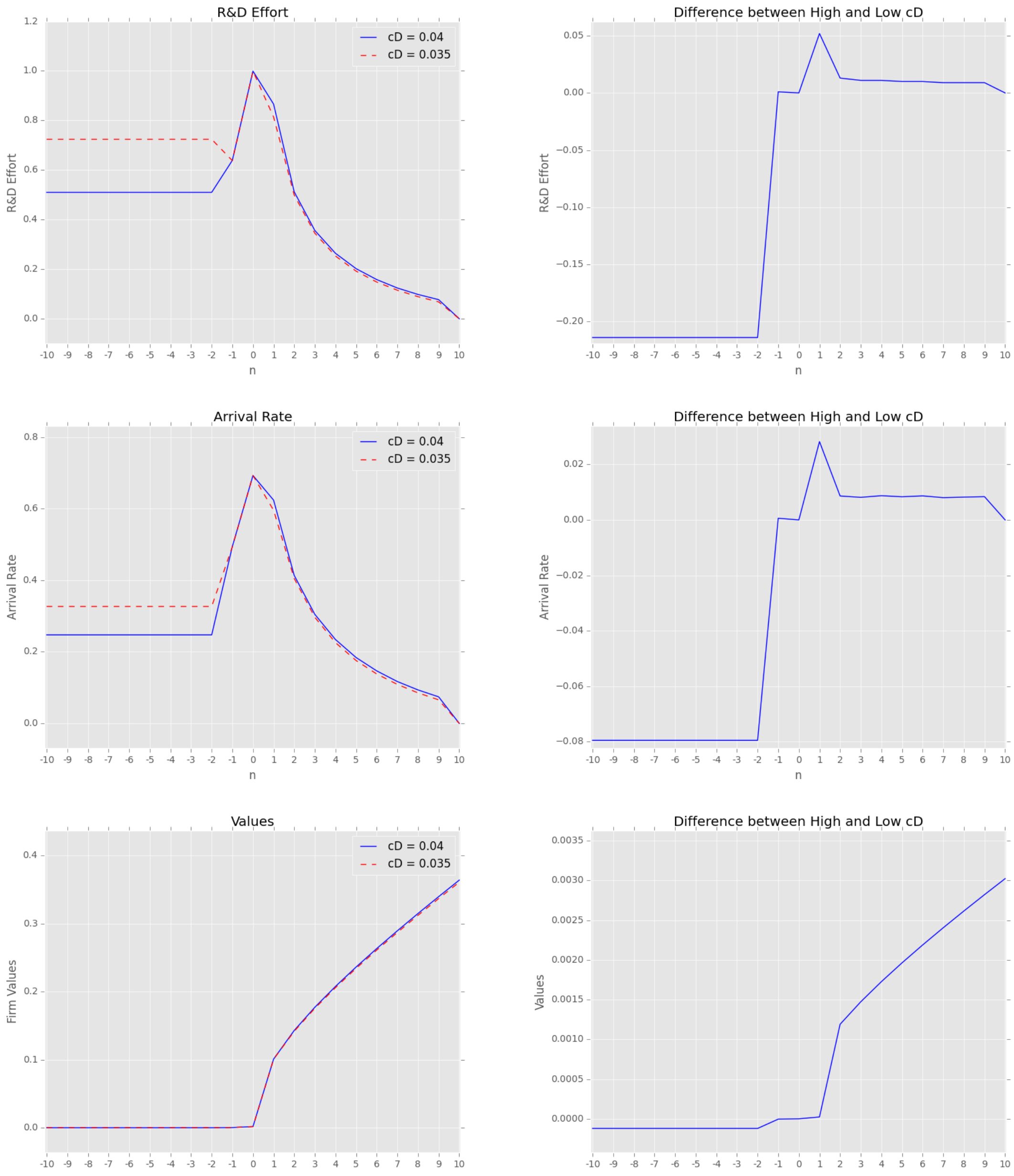

Figure 2: Show the effects a small decrease in the cost per unit of effort for the crazy innovation technology, c^C.
Parameter Values:
$h(x) = \alpha \ln(x+1)$, $\beta = .96$, $\pi_n = 3\ln(\sqrt{n}+1) + .5n$, $c^S_{-n} = e^{\sqrt{n}-2}/6$, $c^C = .045$, $c^{C_{alt}} = .04$, $\alpha = .6$.

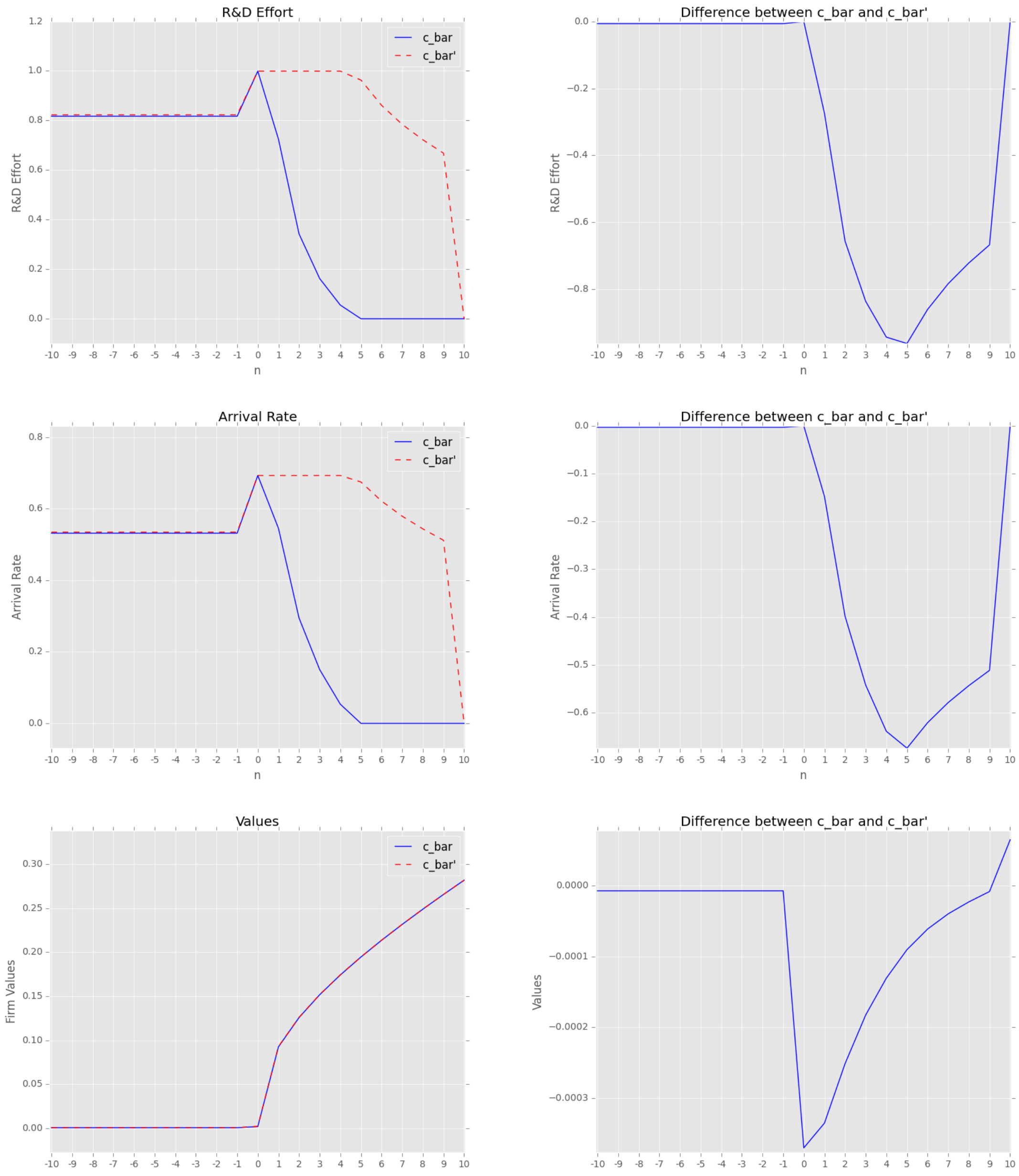

Figure 3: Show the effects of a decrease in the cost per unit effort for the leader when it uses the standard innovation technology, $\bar{c}$.
Parameter Values:
$h(x) = \alpha \ln(x+1)$, $\beta = .96$, $\pi_n = 3\ln(\sqrt{n}+1) + .5n$, $c^S_{-n} = e^{\sqrt{n}-2}/7$, $\bar{c}' = e^{-2}/14$, $c^S = .045$, $\alpha = .6$.

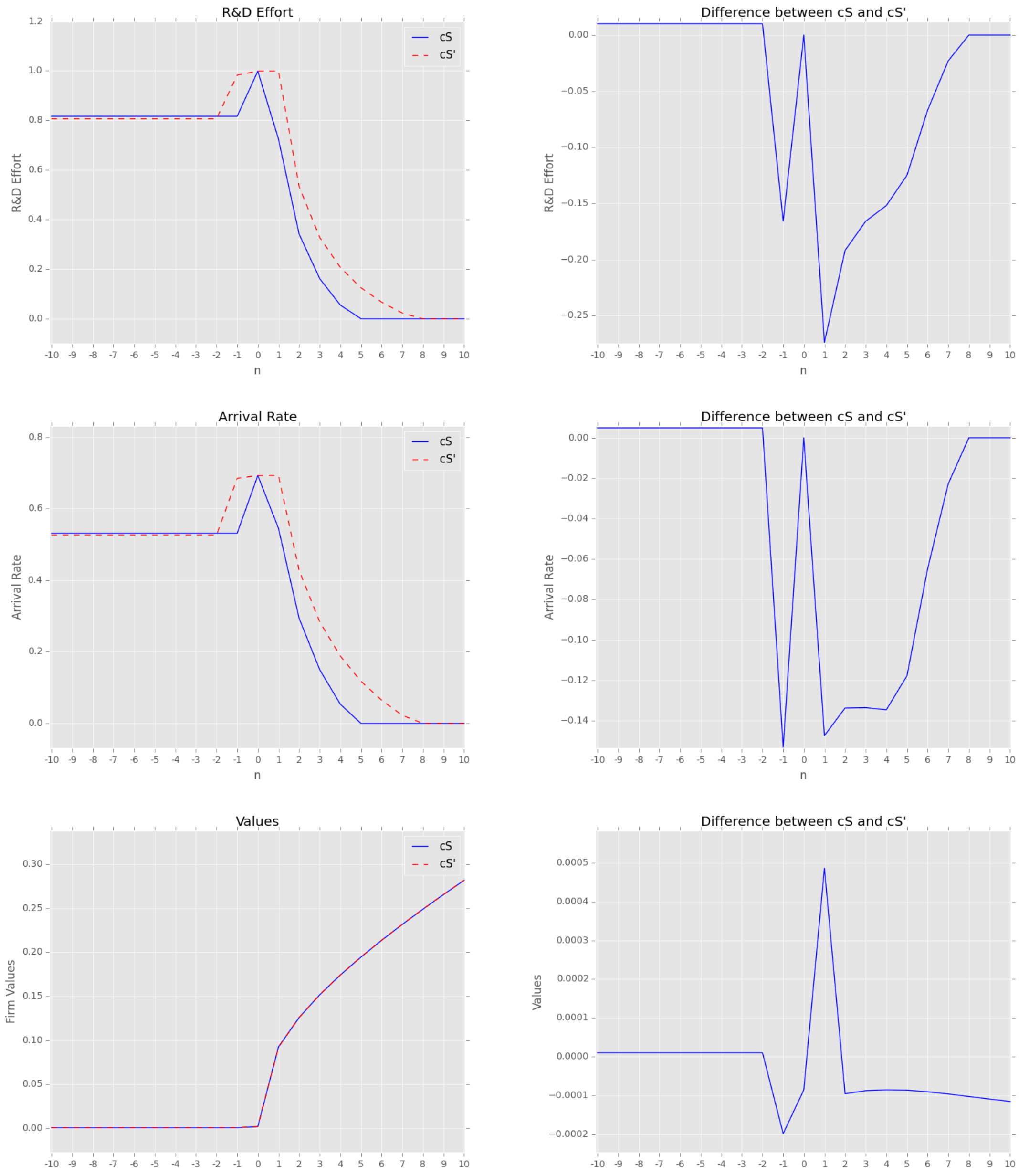

Figure 4: Show the effects of a proportional decrease in the cost per unit of effort for the standard innovation technology, c_n^S. The cost decreases for every n.
Parameter Values:
$h(x) = \alpha \ln(x+1)$, $\beta = .96$, $\pi_n = 3\ln(\sqrt{n}+1) + .5n$, $c_{-n}^S = e^{\sqrt{n}-2}/7$, $c_{-n}^{S'} = e^{\sqrt{n}-2}/8$, $c^C = .045$, $\alpha = .6$.

B Proofs and Derivations

B.1 Derivation of Bellman Equations

As a concrete example, consider a leader that is n steps ahead of its rival. Fix the R&D effort of the other firm, x^*_{-n} and the firm's value in other states. Then the value for the firm at time t is

$$V^S_n(t) = \max_{x^S_n}\{-\bar{c}x^S_n\Delta t + e^{-r(t+\Delta t)\Delta t}[(x^S_n\Delta t + o(\Delta t))V^S_{n+1} + (x^*_{-n}\Delta t + o(\Delta t))V_{-1} \\ + (1 - x^S_n\Delta t - x^*_{-n}\Delta t - o(\Delta t))V^S_n(t+\Delta t)]\}.$$

Subtract $e^{-r(t+\Delta t)\Delta t}V^S_n(t+\Delta t)$ from both sides, divide by Δt, and take the limit as $\Delta t \to 0$. This gives

$$rV^S_n(t) = \max_{x^S_n}\{-\bar{c}x^S_n + x^S_n[V^S_{n+1} - V^S_n(t)] + x^*_{-n}[V_{-1} - V^S_n(t)] + \dot{V}^S_n(t)\}.$$

Now notice that $\dot{V}_0(t) = 0$ and therefore that the firm's value does not depend on t. This gives Equation 4. Equations 5, 6, and 7 are derived similarly.

B.2 Proof of Proposition 1

First, establish that $\exists\ \hat{n}$ such that $x^*_n = 0\ \forall\ n \geq \hat{n}$. To see this, first note that $V^S_n \leq \frac{\pi}{r}$. If $\{V^S_n\}^\infty_n$ is eventually increasing, then the fact that it is also bounded above means that it is eventually concave. Therefore $V^S_{n+1} - V^S_n \to 0$ as $n \to \infty$, and hence for some finite n it must be the case that $x^*_n = \max\{h'^{-1}(\frac{\bar{c}}{V^S_{n+1} - V^S_n}), 0\} = 0$. If instead $\{V^S_n\}^\infty_n$ is not eventually increasing, then for some $\bar{\bar{n}}$, $V^S_{\bar{\bar{n}}+1} < V^S_{\bar{\bar{n}}}$, and hence $x^*_{\bar{\bar{n}}} = 0$. In either case $\exists\ \hat{n}$ such that $x^*_n = 0\ \forall\ n \geq \hat{n}$. Thus, the state space is $\{0, 1, ..., \hat{n}\}$. Note that this fact holds without establishing any other properties of $\{V^S_n\}^\infty_{n=-\infty}$ or $\{x^*_n\}^\infty_{n=-\infty}$.

Now notice that the sets of players and states are finite, the action space is $x_n \in [0, \bar{x}]\ \forall\ n$ and therefore payoffs and transition probabilities $h(x_i) \in [0, \bar{h}]$ has closed graph and is non-empty and convex. Then existence of the equilibrium follows from a straightforward application of Kakutani's fixed point theorem. See, for example, Federgruen (1978).

I next establish each property in turn.

a) First, note that by setting $x_n^* = 0 \; \forall \; n$, the firm guarantees $V_n^S \geq 0 \; \forall \; n$. Since flow profits $\pi_n \leq \bar{\pi} \; \forall \; n$, $V_n^S \leq \frac{\bar{\pi}}{r} \; \forall \; n$, so $\{V_n^S\}_{n=-\infty}^{\infty}$ is bounded below by zero and above by $\frac{\bar{\pi}}{r}$.

b) Suppose by way of contradiction that $V_0^S \geq V_1^S$. Then $x_0^* = \max\{h'^{-1}(\frac{\bar{c}}{V_1^S - V_0^S}), 0\} = 0$. By symmetry of equilibrium, rival's $x_{-0}^* = 0$, so

$$rV_0^S = h(0)[V_1^S - V_0^S] + h(0)[V_{-1}^S - V_0^S] - \bar{c} * 0 = 0,$$

and because $V_0^S \geq V_1^S$ and $V_1^S \geq 0$ by *(a)*, this means that $V_1^S = 0$ as well. Then $x_{-1}^* = \max\{h'^{-1}(\frac{c_{-1}^S}{V_1^S - V_{-1}^S}), 0\} = 0$. But then

$$rV_1^S = h(x_1^*)[V_2^S - V_1^S] + h(0)[V_{-1}^S - V_1^S] + \pi_1 - \bar{c}x_1^*.$$

If $x_1^* > 0$, then it must be that $h(x_1^*)[V_2^S - V_1^S] - \bar{c}x_1^* \geq 0$; otherwise, $x_1^* = 0$. In either case, $V_1^S \geq \frac{\pi_1}{r} > 0$, a contradiction.

c) Suppose by way of contradiction that $V_{-n}^S > V_m^S$. Note that

$$\begin{aligned} r[V_m^S - V_{-n}^S] =& h(x_m^*)[V_{m+1}^S - V_m^S] + h(x_{-m}^*)[V_{-1}^S - V_m^S] + \pi_m - \bar{c}x_m^* \\ & - h(x_{-n}^*)[V_1^S - V_{-n}^S] - h(x_n^*)[V_{-(n+1)}^S - V_{-n}^S] - c_{-n}^S x_{-n}^*. \end{aligned}$$

Need to check 2 cases:

1) $V_{m+1}^S \geq V_m^S$

2) $V_m^S > V_{m+1}^S$

1) By the optimality of x_{-n}^*,

$$\begin{aligned} r[V_1^S - V_{-n}^S] \leq& h(x_m^*)[V_{m+1}^S - V_m^S] + h(x_{-m}^*)[V_{-1}^S - V_m^S] + \pi_m - \bar{c}x_m^* \\ & - h(x_m^*)[V_1^S - V_{-n}^S] - h(x_n^*)[V_{-(n+1)}^S - V_{-n}^S] - c_{-n}^S x_m^* \end{aligned}$$

and by the optimality of x^*_m,

$$r[V^S_m - V^S_{-n}] \geq h(x^*_{-n})[V^S_{m+1} - V^S_m] + h(x^*_{-m})[V^S_{-1} - V^S_m] + \pi_m - \bar{c}x^*_{-n}$$
$$- h(x^*_{-n})[V^S_1 - V^S_{-n}] - h(x^*_n)[V^S_{-(n+1)} - V^S_{-n}] - c^S_{-n}x^*_{-n}.$$

Combining these two inequalities gives that

$$V^S_m - V^S_{-n} \geq V^S_{m+1} - V^S_m + \frac{x^*_{-n} - x^*_m}{h(x^*_{-n}) - h(x^*_m)}[c^S_{-n} - \bar{c}]$$

where the first term is weakly positive because $V^S_{m+1} \geq V^S_m$ and the second term is strictly positive because $h(\cdot)$ is monotonically increasing in x and $c^S_{-n} > \bar{c}$. So $V^S_m - V^S_{-n} > 0$, a contradiction.

2) Just as before, the optimality of x^*_{-n} gives

$$r[V^S_m - V^S_{-n}] \leq h(x^*_m)[V^S_{m+1} - V^S_m] + h(x^*_{-m})[V^S_{-1} - V^S_m] + \pi_m - \bar{c}x^*_m$$
$$- h(x^*_m)[V^S_1 - V^S_{-n}] - h(x^*_n)[V^S_{-(n+1)} - V^S_{-n}] - c^S_{-n}x^*_m.$$

Note that since $V^S_{m+1} < V^S_m$, $x^*_m = \max\{h'^{-1}(\frac{\bar{c}}{V^S_{m+1} - V^S_m}), 0\} = 0$, and so $x^*_{-n} \geq x^*_m$. This means that

$$r[V^S_m - V^S_{-n}] \geq h(x^*_m)[V^S_{m+1} - V^S_m] + h(x^*_{-m})[V^S_{-1} - V^S_m] + \pi_m - \bar{c}x^*_{-n}$$
$$- h(x^*_{-n})[V^S_m - V^S_{-n}] - h(x^*_n)[V^S_{-(n+1)} - V^S_{-n}] - c^S_{-n}x^*_{-n}.$$

Again combining the inequalities gives

$$-\bar{c}x^*_m - h(x^*_m)[V^S_1 - V^S_{-n}] + c^S_{-n}x^*_m \geq -\bar{c}x^*_{-n} - h(x^*_{-n})[V^S_1 - V^S_{-n}] + c^S_{-n}x^*_{-n}.$$

Finally, plugging in 0 for x^*_m gives

$$V^S_m - V^S_{-n} \geq \frac{x^*_{-n}}{h(x^*_{-n})}[c^S_{-n} - \bar{c}],$$

where the left hand side is strictly positive by the logic above, a contradiction.

d) $x^*_{-(n+1)} \leq x^*_{-n} \implies \frac{c^S_{-(n+1)}}{c^S_{-n}} \geq \frac{V^S_1 - V^S_{-(n+1)}}{V^S_1 - V^S_{-n}}$. To show that effort is indeed decreasing as the follower falls further behind, construct an upper bound for the right hand side and show that Assumption 1 guarantees that the left hand side is greater than this upper bound $\forall\ n < 0$. It is clear that

$$\frac{V^S_1 - V^S_{-(n+1)}}{V^S_1 - V^S_{-n}} \leq \frac{V^S_1}{V^S_1 - V^S_{-n}} \leq \frac{\bar{\pi}}{r(V^S_1 - V^S_{-n})}.$$

From *(c)*, know that $V^S_1 - V^S_{-n}$ is bounded away from zero $\forall\ n \leq 0$. In particular,

$$V^S_1 - V^S_{-n} \geq \min\{\frac{x^*_{-n} - x^*_1}{h(x^*_{-n}) - h(x^*_1)}[c^S_{-n} - \bar{c}], \frac{x^*_{-n}}{h(x^*_{-n})}[c^S_{-n} - \bar{c}]\} \geq \frac{1}{h'(0)}[c^S_{-1} - \bar{c}]$$

Plugging in this expression gives $\frac{V^S_1 - V^S_{-(n+1)}}{V^S_1 - V^S_{-n}} \leq \frac{\bar{\pi} h'(0)}{r[c^S_{-1} - \bar{c}]}$, which is less than $\frac{c^S_{-1}}{\bar{c}}$ by Assumption 1, and by definition of c^S_n, know that $\frac{c^S_{-(n+1)}}{c^S_{-n}} > \frac{c^S_{-1}}{\bar{c}}$. Therefore $x^*_{-(n+1)} \leq x^*_{-n}\ \forall\ -n \leq 0$.

e) Recall that $x^*_{-n} = \max\{h^{'-1}(\frac{c_{-n}}{V^S_1 - V^S_{-n}}), 0\}$. Since the denominator is bounded by $\frac{\bar{\pi}}{r}$, the numerator is increasing and convex, and $h^{'-1}(\cdot)$ is a decreasing function, $\exists$ some $-\check{n}$ such that $h^{'-1}(\frac{c^S_{-n}}{V^S_1 - V^S_{-n}}) \leq h^{'-1}(\frac{c^S_{-n}}{V^S_1}) \leq 0$, and so $\forall\ -n \leq -\check{n}$, $x^*_{-n} = 0$.

f) First, show that $V_{n+1} \geq V_n\ \forall\ n \geq \underline{n}$. Notice that $\forall\ n \geq \check{n}$, $V^S_{-n} = V^S_{-\underline{n}} = 0$. Let $n \geq \underline{n}$ and suppose by way of contradiction that $V^S_n \geq V^S_{n+1}$. Then $x^*_n = \max\{h^{'-1}(\frac{\bar{c}}{V^S_{n+1} - V^S_n}), 0\} = 0$. This means that

$$\begin{aligned} V^S_n &= = \frac{\pi_n}{r} \\ &< \frac{\pi_{n+1}}{r} \\ &\leq \frac{h(x^*_{n+1})V^S_{n+2} + \pi_{n+1} - \bar{c}x^*_{n+1}}{r + h(x^*_{n+1})} = V^S_{n+1}, \end{aligned}$$

a contradiction, where the first inequality comes from the fact that $\pi_n < \pi_{n+1}$ and the second inequality comes from the optimality of x^*_{n+1} and the fact that $x^*_{n+1} \geq 0$. The rest of the proof proceeds by backwards induction. Let $n+1=\check{n}$. Suppose by way of contradiction that $V^S_{\check{n}-1} \geq V^S_{\check{n}}$. Then

$$\begin{aligned}
rV^S_{\check{n}} =& h(x^*_{\check{n}})[V^S_{\check{n}+1} - V^S_{\check{n}}] + h(x^*_{-\check{n}})[V^S_{-1} - V^S_{\check{n}}] + \pi_{\check{n}} - \bar{c}x^*_{\check{n}} \\
\geq& h(x^*_{\check{n}-1})[V^S_{\check{n}+1} - V^S_{\check{n}}] + h(x^*_{-\check{n}})[V^S_{-1} - V^S_{\check{n}}] + \pi_{\check{n}} - \bar{c}x^*_{\check{n}-1} \\
\geq& h(x^*_{\check{n}-1})[V^S_{\check{n}} - V^S_{\check{n}-1}] + h(x^*_{-(\check{n}-1)})[V^S_{-1} - V^S_{\check{n}-1}] + \pi_{\check{n}-1} - \bar{c}x^*_{\check{n}-1} \\
=& rV^S_n,
\end{aligned}$$

a contradiction, where the first inequality comes from the optimality of $x^*_{\check{n}}$ and the second inequality comes from the fact that $V^S_{\check{n}+1} \geq V^S_{\check{n}}$ as shown above, the assumption that $V^S_{\check{n}-1} > V_{\check{n}}$, $\pi_{\check{n}} > \pi_{\check{n}-1}$, the fact that $V^S_{-1} < V^S_{\check{n}-1}$ from *c*, and $x^*_{-\check{n}} \leq x^*_{-(\check{n}-1)}$ from *d*. Now assume by way of contradiction that $V^S_{\check{n}-2} > V^S_{\check{n}-1}$, using the fact that $V^S_{\check{n}-1} \leq V^S_{\check{n}}$. A contradiction is reached by identical logic as before. Continuing this process establishes the result $\forall\ n \geq 0$.

g) Define $\delta_n \equiv V^S_n - V^S_{n-1}$ and suppose by way of contradiction that $\delta_{n+1} - \delta_n > 0$. Then using the optimality of x^*_n and the fact that $x^*_{-n} > x^*_{-(n+1)}\ \forall\ -n < 0$ from *(d)*, have that

$$\begin{aligned}
rV^S_{n+1} =& h(x^*_{n+1})[V^S_{n+2} - V^S_{n+1}] + h(x^*_{-(n+1)})[V^S_{-1} - V^S_{n+1}] + \pi_{n+1} - \bar{c}x^*_{n+1} \\
\leq& h(x^*_{n+1})[V^S_{n+2} - V^S_{n+1}] + \pi_{n+1} - \bar{c}x^*_{n+1}, \\
rV_n \geq& h(x^*_{n+1})[V^S_{n+1} - V^S_n] + h(x^*_{-n})[V^S_{-1} - V^S_n] + \pi_n - \bar{c}x^*_{n+1}, \\
rV^S_n \geq& h(x^*_{n-1})[V^S_{n+1} - V^S_n] + h(x^*_{-n})[V^S_{-1} - V^S_n] + \pi_n - \bar{c}x^*_{n-1},
\end{aligned}$$

and

$$\begin{aligned}
rV^S_{n-1} =& h(x^*_{n-1})[V^S_n - V^S_{n-1}] + h(x^* - (n-1))[V^S_{-1} - V^S_{n-1}] + \pi_{n-1} - \bar{c}x^*_{n-1} \\
\leq& h(x^*_{n-1})[V^S_n - V^S_{-n-1}] + \pi_{n-1} - \bar{c}x^*_{n-1}.
\end{aligned}$$

Taking difference gives

$$r[V^S_{n+1} - V^S_n] \leq h(x^*_{n+1})[V^S_{n+1} + V^S_n - 2V^S_{n+1}] + h(x^*_{-n})[V^S_{-1} - V^S_n] + \pi_{n+1} - \pi_n$$

and

$$r[V^S_n - V^S_{n-1}] \geq h(x^*_{n-1})[V^S_{n+1} + V^S_{n-1} - 2V^S_n] + h(x^*_{-n})[V^S_{-1} - V^S_n] + \pi_n - \pi_{n-1},$$

which, after rewriting, means that

$$r\delta_{n+1} \leq h(x^*_{n+1})[\delta_{n+2} - \delta_{n+1}] + h(x^*_{-n})[V^S_{-1} - V^S_n] + \pi_{n+1} - \pi_n$$

and

$$r\delta_n \geq h(x^*_{n-1})[\delta_{n+1} - \delta_n] + h(x^*_{-n})[V^S_{-1} - V^S_n] + \pi_n - \pi_{n-1}.$$

Again taking differences gives

$$(r + h(x^*_{-n}))[\delta_{n+1} - \delta_n] \leq h(x^*_{n+1})[\delta_{n+2} - \delta_{n+1}] + \pi_{n+1} + \pi_{n-1} - 2\pi_n.$$

Since $\pi_{n+1} + \pi_{n-1} - 2\pi_n < 0$, the assumption that $\delta_{n+1} - \delta_n \geq 0$ means that $\delta_{n+2} - \delta_{n+1} > 0$. Repeating this argument successively means that if $\delta_{n'+1} - \delta_{n'} \geq 0$, then $\delta_{n+1} - \delta_n > 0 \ \forall \ n > n'$. But know $V^S_{n+1} > V^S_n \ \forall \ n \geq 0$ from *(f)* and $\{V^S_n\}^\infty_{n=0}$ converges to $\frac{\bar{\pi}}{r} < \infty$ from *(a)*, thus for sufficiently large n $\delta_n \to 0$ and hence $\delta_{n+1} - \delta_n \leq 0$, a contradiction.

h) By *(g)*, know $V^S_{n+1} - V^S_n < V^S_n - V^S_{n-1}$, and so $\frac{\bar{c}}{V^S_{n+1} - V^S_n} < \frac{\bar{c}}{V^S_n - V^S_{n-1}}$, and since $h'^{-1}(\cdot)$ is a decreasing function, $x^*_{n-1} > x^*_n$ as long as $\max\{h'^{-1}(\frac{\bar{c}}{V^S_n - V^S_{n-1}}), 0\} > 0$. By the monotonicity of V^S_n, $\exists \ \bar{n}$ such that $h'^{-1}(\frac{\bar{c}}{V^S_{\bar{n}} - V^S_{\bar{n}-1}}) > 0$ and $h'^{-1}(\frac{\bar{c}}{V^S_{\bar{n}+1} - V^S_{\bar{n}}}) = 0$, and therefore $x^*_n = 0 \ \forall \ n \geq \bar{n}$.

i) First, as noted in *(f)*, clearly $\forall \ -n \leq -\check{n}$, $V^S_{-n} = V^S_{-\check{n}} = 0$. Now let $-(n+1) = -\check{n}$.

Suppose by way of contradiction that $V^S_{-\check{n}} > V^S_{-(\check{n}-1)}$. Then

$$\begin{aligned}
rV^S_{-(\check{n}-1)} =& h(x^*_{-(\check{n}-1)})[V^S_1 - V^S_{-(\check{n}-1)}] + h(x^*_{\check{n}-1})[V^S_{-\check{n}} - V^S_{-(\check{n}-1)}] - c^S_{-(\check{n}-1)}x^*_{-(\check{n}-1)} \\
\geq & h(x^*_{-\check{n}})[V^S_1 - V^S_{-(\check{n}-1)}] + h(x^*_{\check{n}-1})[V^S_{-\check{n}} - V^S_{-(\check{n}-1)}] - c^S_{-(\check{n}-1)}x^*_{-\check{n}} \\
\geq & h(x^*_{-\check{n}})[V^S_1 - V^S_{-\check{n}}] + h(x^*_{\check{n}})[V^S_{-\check{n}} - V^S_{-\check{n}}] - c^S_{-\check{n}}x^*_{-(\check{n}-1)} \\
=& rV^S_{\check{n}},
\end{aligned}$$

a contradiction, where the first inequality comes from the optimality of $x^*_{-(\check{n}-1)}$ and the second inequality comes from the fact that $V^S_{-(\check{n}+1)} = V^S_{-\check{n}}$, the assumption that $V^S_{-\check{n}} > V^S_{-(\check{n}-1)}$, and the fact that $c^S_{-\check{n}} > c^S_{-(\check{n}-1)}$. Now proceed by backwards induction, using the fact that $V^S_{-(\check{n}-1)} \geq V^S_{-\check{n}}$. Suppose by way of contradiction that $V^S_{-(\check{n}-1)} > V^S_{-(\check{n}-2)}$. Identical steps to those above give a contradiction. Likewise $\forall \ -n \leq 0$. Therefore, $V^S_{-n} \geq V^S_{-n} \forall \ -n \leq 0$.

□

B.3 Proof of Proposition 2

Consider $-n \leq -\hat{n}$. Then,

$$V^C_{-n} = \frac{-c^C x^*_{-n} + h(x^*_{-n})V^S_1}{r + h(x^*_{-n})}.$$

None of these parameters depend on n, so the optimal choice $x^*_{-n} = x^C$ and then $V^C_{-n} = V^C$. Now let $-n = -(\hat{n}-1)$. Suppose by way of contradiction that $V^C_{-(\hat{n}+1)} > V^C$. Then

$$-c^C x^*_{-(\hat{n}-1)} + \alpha h(x^*_{-(\hat{n}-1)})[V^S_1 - V^C_{-(\hat{n}-1)}] + h(x^*_{\hat{n}-1})[V^C - V^C_{-(\hat{n}-1)}] > -c^C x^*_{-\hat{n}} + \alpha h(x^*_{-\hat{n}})[V^S_1 - V^C].$$

But

$$\begin{aligned}
-c^C x^*_{-\hat{n}} + \alpha h(x^*_{-\hat{n}})[V^S_1 - V^C] \geq & - c^C x^*_{-(\hat{n}-1)} + \alpha h(x^*_{-(\hat{n}-1)})[V^S_1 - V^C] \\
\geq & - c^C x^*_{-(\hat{n}-1)} + \alpha h(x^*_{-(\hat{n}-1)})[V^S_1 - V^C_{-(\hat{n}-1)}] \\
> & - c^C x^*_{-(\hat{n}-1)} + \alpha h(x^*_{-(\hat{n}-1)})[V^S_1 - V^C] + h(x^*_{\hat{n}-1})[V^C - V^C_{-(\hat{n}-1)}],
\end{aligned}$$

a contradiction, where the first inequality comes from the optimality of $x^*_{-\hat{n}}$ and the second and third inequalities use the fact that $V^C_{-(\hat{n}+1)} > V^C$.

Now suppose by way of contradiction that $V^C_{-(\hat{n}+1)} > V^C$. Then

$$-c^C x^*_{-(\hat{n}-1)} + \alpha h(x^*_{-(\hat{n}-1)})[V^S_1 - V^C_{-(\hat{n}-1)}] + h(x^*_{\hat{n}-1})[V^C - V^C_{-(\hat{n}-1)}] < -c^C x^*_{-\hat{n}} + \alpha h(x^*_{-\hat{n}})[V^S_1 - V^C].$$

But

$$\begin{aligned}
-c^C x^*_{-(\hat{n}-1)} + \alpha h(x^*_{-(\hat{n}-1)})[V^S_1 - V^C_{-(\hat{n}-1)}] + h(x^*_{\hat{n}-1})[V^C - V^C_{-(\hat{n}-1)}]& \\
\geq - c^C x^*_{-\hat{n}} + \alpha h(x^*_{-\hat{n}})[V^S_1 - V^C_{-(\hat{n}-1)}] + h(x^*_{\hat{n}-1})[V^C - V^C_{-(\hat{n}-1)}]& \\
\geq - c^C x^*_{-\hat{n}} + \alpha h(x^*_{-\hat{n}})[V^S_1 - V^C_{-(\hat{n}-1)}] + h(x^*_{\hat{n}-1})[V^C - V^C]& \\
> - c^C x^*_{-\hat{n}} + \alpha h(x^*_{-\hat{n}})[V^S_1 - V^C],&
\end{aligned}$$

a contradiction, where the first inequality comes from the optimality of $x^*_{-(\hat{n}-1)}$ and the second and third inequalities use the fact that $V^C < V^C_{-(\hat{n}-1)}$.

Thus $V^C = V^C_{-(\hat{n}-1)}$. Identical reasoning establishes that $V^C = V^C_{-n}\ \forall -n \in [-(\hat{n}-1), 0)$.

The remainder of the proof follows nearly identically from Acemoglu (2009) and Acemoglu and Akcigit (2012). I reproduce the arguments here. The fact that $V^S_1 > V^S_0 > V^C$ follows identical steps to the proof of parts *(b)* and *(c)* of Proposition 1.

To show that $V^S_{n+1} > V^S_n\ \forall\ n > 0$, suppose by way of contradiction that $V^S_{n+1} \leq V^S_n$. Then $x^*_n = 0$. Note that

$$\begin{aligned}
rV^S_{n+1} =& \pi_{n+1} - \bar{c}x^{S*}_{n+1} + h(x^{S*}_{n+1})[V^S_{n+2} - V^S_{n+1}] + h(x^C)[V^S_{-1} - V^S_{n+1}] \\
\geq& \pi_{n+1} - \bar{c}x^{S*}_{n} + h(x^{S*}_{n})[V^S_{n+2} - V^S_{n+1}] + h(x^C)[V^S_{-1} - V^S_{n+1}] \\
=& \pi_{n+1} + h(x^C)[V^S_{-1} - V^S_{n+1}],
\end{aligned}$$

where there inequality comes from the optimality of x^{S*}_{n+1}. Then using the assumption that $V^S_{n+1} \leq V^S_n$ gives

$$\pi_n + h(x^C)[V^S_{-1} - V^S_n] \geq \pi_{n+1} + h(x^C)[V^S_{-1} - V^S_{n+1}].$$

Since $\pi_{n+1} > \pi_n$, this inequality means that $V_{n+1}^S > V_n^S$, a contradiction.

To show that $x_{n+1}^{S*} \leq x_n^{S*}\ \forall\ n > 0$, first define

$$\delta_n \equiv V_{n+1}^S - V_n^S.$$

Also define $\bar{r} \equiv r + h(x^C)$. Rewrite V_n^S as

$$\bar{r}V_n^S = \pi_n - \bar{c}x_n^{S*} + h(x_n^{S*})[V_{n+1}^S - V_n^S] + h(x^C)V_{-1}^S.$$

The optimality of x_n^{S*} gives the following system of inequalities:

$$\begin{aligned}
\bar{r}V_{n+1}^S &= \pi_{n+1} - \bar{c}x_{n+1}^{S*} + h(x_{n+1}^{S*})[V_{n+2}^S - V_{n+1}^S] + h(x^C)V_{-1}^S \\
\bar{r}V_n^S &\geq \pi_n - \bar{c}x_{n+1}^{S*} + h(x_{n+1}^{S*})[V_{n+1}^S - V_n^S] + h(x^C)V_{-1}^S \\
\bar{r}V_n^S &\geq \pi_n - \bar{c}x_{n-1}^{S*} + h(x_{n-1}^{S*})[V_{n+1}^S - V_n^S] + h(x^C)V_{-1}^S \\
\bar{r}V_{n-1}^S &= \pi_{n-1} - \bar{c}x_{n-1}^{S*} + h(x_{n-1}^{S*})[V_n^S - V_{n-1}^S] + h(x^C)V_{-1}^S.
\end{aligned}$$

Taking differences and using the definition of δ_n gives

$$\begin{aligned}
\bar{r}\delta_{n+1} &\leq \pi_{n+1} - \pi_n + h(x_{n+1}^{S*}[\delta_{n+2} - \delta_{n+1}] \\
\bar{r}\delta_n &\geq \pi_n - \pi_{n+1} + h(x_{n-1}^{S*}[\delta_{n+1} - \delta_n],
\end{aligned}$$

and therefore

$$(\bar{r} + h(x_{n-1}^{S*})(\delta_{n+1} - \delta_n) \leq \pi_{n+1} + \pi_{n-1} - 2\pi_n + h(x_{n+1}^{S*})[\delta_{n+2} - \delta_{n+1}].$$

Suppose by way of contradiction that $\delta_{n+1} - \delta_n \geq 0$. But from the previous inequality, this means that $\delta_{n+2} - \delta_{n+1} > 0$ since $\pi_{n+1} + \pi_{n-1} - 2\pi_n > 0$. Repeating the argument successively gives that if $\delta_{n'+1} - \delta_{n'} > 0$, then $\delta_{n+1} - \delta_n > 0\ \forall\ n \geq n'$. But the previous part of this proof established that $\{V_n^S\}_{n=0}^{\infty}$ is strictly increasing and converges to a finite constant, so $\delta_{n+1} - \delta_n < 0$ for sufficiently large n, a contradiction. Thus, $x_{n+1}^{S*} \leq x_n^{S*}$, and the inequality clearly must be strict if $x_n^{S*} > 0$, which occurs for all $n < \hat{n}$. $\square$

B.4 Proof of Lemma 1

Suppose by way of contradiction that $\exists\ \tilde{n}$ such that $V^C_{-\tilde{n}} \geq V^S_{-\tilde{n}}$ and $V^S_{-(\tilde{n}+1)} > V^C_{-(\tilde{n}+1)}$. This means that either $V^S_{-\tilde{n}} < V^S_{-(\tilde{n}+1)}$ or $V^C_{-\tilde{n}} > V^C_{-(\tilde{n}+1)}$.

First, suppose by way of contradiction that $V^S_{-\tilde{n}} < V^S_{-(\tilde{n}+1)}$. But

$$\begin{aligned}
rV^S_{-\tilde{n}} =& -c^S_{-\tilde{n}}x^{S*}_{-\tilde{n}} + h(x^{S*}_{-\tilde{n}})[V^S_1 - V^S_{-\tilde{n}}] + h(x^{S*}_{\tilde{n}}[V^S_{-(\tilde{n}+1)} - V^S_{-\tilde{n}}] \\
\geq& -c^S_{-\tilde{n}}x^{S*}_{-(\tilde{n}+1)} + h(x^{S*}_{-(\tilde{n}+1)})[V^S_1 - V^S_{-\tilde{n}}] + h(x^{S*}_{\tilde{n}}[V^S_{-(\tilde{n}+1)} - V^S_{-\tilde{n}}] \\
\geq& -c^S_{-(\tilde{n}+1)}x^{S*}_{-(\tilde{n}+1)} + h(x^{S*}_{-(\tilde{n}+1)})[V^S_1 - V^S_{-(\tilde{n}+1)}] + h(x^{S*}_{\tilde{n}}[V^S_{-(\tilde{n}+1)} - V^S_{-(\tilde{n}+1)}] \\
\geq& -c^S_{-(\tilde{n}+1)}x^{S*}_{-(\tilde{n}+1)} + h(x^{S*}_{-(\tilde{n}+1)})[V^S_1 - V^S_{-(\tilde{n}+1)}] + h(x^{S*}_{(\tilde{n}+1)}[V^S_{-(\tilde{n}+2)} - V^S_{-(\tilde{n}+1)}] \\
=& rV^S_{-(\tilde{n}+1)},
\end{aligned}$$

a contradiction, where the first inequality comes from the optimality of $x^{S*}_{-\tilde{n}}$; the second inequality comes from the assumption that $V^S_{-\tilde{n}} < V^S_{-(\tilde{n}+1)}$ and the fact that $c^S_{-\tilde{n}} < c^S_{-(\tilde{n}+1)}$; the third inequality comes from parts (h) and (i) of Proposition 1, which hold when the leader and follower both work on only the standard innovation.

Thus, it must be the case that $V^C_{-\tilde{n}} > V^C_{-(\tilde{n}+1)}$. But if the follower uses $C\ \forall n > \tilde{n}$, it guarantees itself V^C forever, as proven in Proposition 2. Suppose by way of contradiction that $V^C_{-\tilde{n}} > V^C$. But

$$\begin{aligned}
rV^C =& -c^C x^{R*} + \alpha h(x^{R*})[V^S_1 - V^C] \\
\geq& -c^C x^{R*}_{-\tilde{n}} + \alpha h(x^{R*}_{-\tilde{n}}[V^S_1 - V^C] \\
\geq& -c^C x^{R*}_{-\tilde{n}} + \alpha h(x^{R*}_{-\tilde{n}}[V^S_1 - V^C] + h(x^{S*}_{\tilde{n}}[V^S_{-(\tilde{n}+1)} - V^C_{-\tilde{n}}] \\
\geq& -c^C x^{R*}_{-\tilde{n}} + \alpha h(x^{R*}_{-\tilde{n}}[V^S_1 - V^C_{-\tilde{n}}] + h(x^{S*}_{\tilde{n}}[V^S_{-(\tilde{n}+1)} - V^C_{-\tilde{n}}] \\
=& rV^C_{-\tilde{n}},
\end{aligned}$$

a contradiction, where the first inequality comes from the optimality of x^{R*}, the second inequality comes from the fact that $V^S_{-(\tilde{n}+1)} \leq V^S_{-\tilde{n}} \leq V^C_{-\tilde{n}}$, and the third inequality comes from the assumption that $V^C_{-\tilde{n}} > V^C$.

So, $V^C \geq V^C_{-\tilde{n}}$ and therefore the follower can guarantee itself higher lifetime expected value by never switching back to the standard innovation after $-\tilde{n}$. $\square$

B.5 Proof of Lemma 2

First note that $x^C = h'^{-1}(\frac{c^C}{\alpha[V_1^S - V^C]})$ and $x^*_{-\tilde{n}} = h'^{-1}(\frac{c^S_{-\tilde{n}}}{V_1^S - V^S_{-\tilde{n}}})$.

a) If $x^C > x^*_{-\tilde{n}}$, then by rearranging the above expressions,

$$\frac{V_1^S - V^C}{V_1^S - V^S_{-\tilde{n}}} > \frac{c^C}{\alpha c^S_{-\tilde{n}}}.$$

I derive a sufficient condition to ensure that this inequality holds. In particular, derive a lower bound for the left-hand side. Clearly $V_1^S - V^S_{-\tilde{n}} \leq V_1^S \leq \frac{\bar{\pi}}{r}$. Need to show that $V_1^S - V^C$ is bounded strictly away from zero. Note that

$$\begin{aligned} r[V_1^S - V^C] \leq & h(x_1^*)[V_2^S - V_1^S] + h(x^*_{-1})[V^S_{-1} - V_1^S] + \pi_1 - \bar{c}x_1^* \\ & - \alpha h(x_1^*)[V_1^S - V^C] + c^C x_1^* \end{aligned}$$

and

$$\begin{aligned} r[V_1^S - V^C] \geq & h(x^C)[V_2^S - V_1^S] + h(x^*_{-1})[V^S_{-1} - V_1^S] + \pi_1 - \bar{c}x^C \\ & - \alpha h(x^C)[V_1^S - V^C] + c^C x^C. \end{aligned}$$

Combining these and simplifying gives

$$\begin{aligned} V_1^S - V^C \geq & \frac{1}{\alpha}[V_2^S - V_1^S] + \frac{x^C - x_1^*}{\alpha[h(x^C) - h(x_1^*)]}[c^C - \bar{c}] \\ \geq & \frac{1}{h'(0)}[c^C - \bar{c}] \end{aligned}$$

using arguments identical to Prop. 1 part (c). Hence, if

$$\frac{c^C}{c^S_{-\tilde{n}}} \leq \frac{\alpha r[c^C - \bar{c}]}{h'(0)\bar{\pi}}$$

then it must be true that $x^C \geq x^*_{-\tilde{n}}$.

b) Likewise, $x^*_{-\tilde{n}} \geq x^C$ means that

$$\frac{c^C}{\alpha c^S_{-\tilde{n}}} \geq \frac{V_1^S - V^C}{V_1^S - V^S_{-\tilde{n}}}.$$

I derive a sufficient condition so that this inequality is true. Clearly $V_1^S - V^C \leq V_1^S \leq \frac{\bar{\pi}}{r}$. From part ($d$) of Prop. 1, know that $V_1^S - V^S_{-\tilde{n}} \geq \frac{1}{h'(0)}[c^S_{-1} - \bar{c}]$. So, if

$$\frac{c^D}{c^S_{-\tilde{n}}} \geq \frac{\alpha \bar{\pi} h'(0)}{r[c^S_{-1} - \bar{c}]}$$

then it must be true that $x^*_{-\tilde{n}} \geq x^C$.

c) If either of these sufficient conditions does not hold, then it is ambiguous whether x^C or $x^*_{-\tilde{n}}$ is larger. However, it must be shown that these sufficient conditions do not overlap, that is,

$$\frac{\alpha \bar{\pi} h'(0)}{r[c^S_{-1} - \bar{c}]} > \frac{\alpha r[c^C - \bar{c}]}{h'(0)\bar{\pi}}.$$

Rearranging, this gives that

$$(\bar{\pi} h'(0))^2 > r^2[c^C - \bar{c}] \cdot [c^S_{-1} - c^C].$$

Clearly the left hand side is positive. The right hand side is negative by the assumption that $c^C > c^S_{-1}$.

$\square$