

Patent Protection, Takeovers, and Startup Innovation: A Dynamic Approach

Andreas Panagopoulos* In-Uck Park†

Working Paper‡

August 14, 2008

Abstract

The impact of IP protection on the innovation incentives of startup firms is examined in a dynamic model where an incumbent faces a sequence of potential startups and the incumbent's chance of winning an infringement lawsuit increases with the size of its patent portfolio. It is shown that takeover deals generate extra benefits for the incumbent via its enhanced future bargaining positions, a part of which accrues to the current startup as an increased bargaining share. This increased bargaining share can be large enough to justify the startup's innovation activity that would not have taken place otherwise. This effect may be greatest under moderate levels of IP protection, because the increase in the bargaining share, being proportional to the marginal benefits brought by the last patent added to the portfolio, would be too small if the protection was too weak while it would taper off too quickly if the protection was excessive

Keywords: Patent litigation, takeovers, patent portfolios.

JEL No: O31, O34, L21, L24, K0

*University of Lancaster, a.panagopoulos@lancaster.ac.uk

†University of Bristol, i.park@bristol.ac.uk

‡We would like to thank Philippe Aghion, Jim Bessen, Vincenzo Denicolò, Stuart Graham, Dietmar Harhoff, Gerard Llobet, Eric Maskin, Manolis Petrakis, Bob Rothschild, and the participants at the EPIP conference and the ZEW conference on the economics of innovation and patenting for their helpful comments and suggestions. All errors remain our responsibility. This research is funded by the ESRC small grants program.

1 Introduction

It is widely acknowledged that the creation in 1982 of the Central Appellate Court for the Federal Circuit (CAFC), a centralized patent court in the US that lacks an EU equivalent, strengthened intellectual property (IP) rights. As a consequence, firms became to rely more heavily on patents as a means to protect their IP, a shift from prior practice that has led to an upsurge in patenting and has contributed to the creation of large patent portfolios by dominant US firms. Proponents of this change view stronger IP rights as having provided a fertile ground on which US technological superiority has spawned, contrary to the situation in the EU. Critics, on the other hand, acknowledge a drop in patent quality and frequent overlapping patents that have considerably increased costly patent litigation both in time and money. In addressing these concerns policy makers have introduced the Patent Reform Act of 2007, whose main role is to reduce patent litigation.

Following this trend economists recognized the need to reassess the rationale for patents given that new innovations increasingly tend to be cumulative/sequential, building upon previously-patented prior art. In such environments, due to the inevitable conflict in providing incentives to current and future innovators, the dynamic effects of patent systems can be fundamentally different from the conventional wisdom, as forcefully argued recently by Bessen and Maskin (2007) in a dynamic game between rival innovators, and by Hopenhayn *et al.* (2006) from the social planner's optimal patent design perspective. The current paper contributes to this debate by furthering our understanding of the impact of different levels of IP protection on the long-run innovation dynamics of startup firms. This shift in focus from established patentees to startup innovators is in recognition of the extent to which the latter's innovativeness has shaped modern hi-tech industries, and of their frequent inability to defend their IP rights against dominant rivals (owing to the considerable legal fees and the uncertainty surrounding patent litigation).

In the US where startups are most active, since the foundation of CAFC there has been a steady increase of takeover activities (even after discounting the surge of such activities shortly before 2001, that is attributable to the internet bubble) as shown in Figures 1-2, where established firms ventured to acquire smaller ones. True as it may be that the reasons behind such takeovers are diverse, many were driven by the desire to expand a firm's technological horizons via the purchase of innovative startup firms.¹

¹Restricting the argument to a few prominent examples, in the 1984 to 2001 period G.E. and Siemens acquired more than 110 and 170 firms respectively; see Dessyllas and Hughes (2005).

This drive, coined by H. Chesbrough as *open innovation*, gained prominence among firms such as Intel and Cisco; see Chesbrough (2003). The prospects of such takeovers serve as a secondary market for ideas, in addition to the NASDAQ, offering entrepreneurial entrants an extra option for reaping the benefits of their inventions. This co-evolution of the court's attitude and takeover activities is consistent with the intuition behind this short paper.

Operating under the constant threat of infringement allegations from an incumbent firm, a startup's innovation incentive is influenced by the level of IP protection that the legal system would provide, in the form of upholding the incumbent's patent infringement allegation against the startup. Along these lines, in a model where an incumbent faces a sequence of potential startups and the incumbent's chance of winning an infringement lawsuit increases with the size of its patent portfolio, we show that takeover deals (out of court settlements) generate extra benefits for the incumbent via its enhanced bargaining positions in future settlement deals, and a part of such extra benefits accrues to the current startup as an increased bargaining share. This increased bargaining share may actually be large enough to justify the startup's innovation activity that would not have taken place otherwise. This effect may be greatest under moderate levels of IP protection, because the increase in the bargaining share, being proportional to the marginal benefits brought by the last patent added to the portfolio, would be too small if the protection was too weak while it would taper off too quickly if the protection was excessive.

Implicit in the argument above appears to be a presumption we adopt, namely, that an incumbent firm has accumulated a patent portfolio on a set of technologies that are interrelated and advance on a common theme (having a central idea as a backbone) and hence, form a unified technological territory. Although in reality large firms' patent portfolios tend to range across many and frequently diverging technologies, for the purpose of this paper a patent portfolio means a subset of a firm's patents that are on a specific technological terrain.² Thus, the larger is a firm's patent portfolio, the better described, entrenched, and protected becomes its technological territory, enhancing the firm's ability to prevail in court over infringement disputes concerning newer ideas that are perceived as

²Chemical firms frequently form such portfolios. For example, upon the invention of NYLON DuPont patented all the chemical formulae bearing resemblance to its core technology. More recently, with regard to the wireless market contended by tech giants such as Apple Inc., *Forbes* wrote "The end game: futuristic gizmos controlled by gestures that are tied wirelessly to the world around them, *protected by a broad portfolio of patents, . . .*" (emphasis added); see http://www.forbes.com/2008/04/30/apple-iphone-3g-tech-wire-cx_bc_0501apple.html?partner=yahootix.

having built upon the same technological terrain. This association between patent portfolios and the ability to protect an individual patent is supported by empirical findings, e.g., Lanjouw and Schankerman (2004) explained in Section 2.

In particular, in hi-tech industries where technology is complex and cumulative, a new innovation is likely to infringe on existing patents. Therefore, an entrant firm is likely to face the threat of litigation from a competing incumbent patent holder of a significantly larger size. Commentators argue that when IP protection is limited, and the courts have lenient views regarding alleged infringement, an entrant will abstain from innovating because its innovation will not be adequately protected from the incumbent. However, the flip side of this logic also suggests that if IP protection is strong, and the courts take a tough stance on alleged infringement, the fear of a difficult litigation battle would equally diminish the entrant's incentives to innovate.

As will be shown later, if IP protection is moderate, allowing the courts a balanced approach towards alleged infringement, the benefit of a takeover for the incumbent goes beyond commercializing the new innovation. This is because the incumbent capitalizes on the enhanced bargaining position that the current takeover will bring in all potential future deals by incorporating the current startup's patented ideas to its own patent portfolio, which allows it to better barricade its technological territory. Since this prospect of future surplus for the incumbent hinges on the current takeover, a part of the surplus accrues to the current startup, enlarging its bargaining share. We show that this dynamic effect of moderate IP protection can motivate the startups' innovation activities that would not take place without it, and as a consequence, increase the social welfare. We emphasize, however, that for maximum effect the level of IP protection should be selected carefully because excessive IP protection would accumulate the incumbent's bargaining power too quickly, killing off the innovation incentives for startups prematurely.

The existing literature on cumulative innovations includes Scotchmer (1996), Green and Scotchmer (1995), and Chang (1995), that focus on a single follow-on innovation; and O'Donoghue, *et al.* (1998), Hopenhayn, *et al.* (2006), and Bessen and Maskin (2007), on multiple sequential innovations. The main feature that differentiates our paper is that we explicitly model the uncertain nature of court rulings in infringement suits, based on how dissimilar firms (incumbents *vs.* startups), having amassed patent portfolios to indifferent depths, defend their IP rights. Our approach shares some key features of, and is complementary to, Bessen and Maskin (2007) who examine the dynamic interaction between equally dominant firms, with an important distinction that our results do not

need “complementarities” between innovation efforts that appear to be essential in their analysis.

Defining boundaries of intellectual properties is inherently imperfect and as a result, disputes are inevitable and the court plays an active role in the way the patent system operates.³ The effects of patent litigation have been studied by Meurer (1989), Choi (1998), Aoki and Hu (1999), Crampes and Langinier (2002), and Llobet (2003), however these authors mostly dealt with a single patent to protect and thus, their foci of analysis differ from ours: In a model where the incumbent’s patent portfolio can evolve over time, we examine the feedback effect that the prospect of such evolution may have on the startup’s innovation incentives, and discuss the long-run welfare implications thereof. Lastly, even though our argument is delivered through a functioning patent system it is nevertheless related to the message of a growing literature that argues against the current IP system because of its excessive protection; see Boldrin and Levine (2002).

Section 2 is a brief review of the patent system that provides some background for our analysis. Section 3 presents a static model as a benchmark. Section 4 provides the main analysis of the dynamic model and characterizes the unique equilibrium. Section 5 discusses some implications of our findings using simulations. Section 6 contains some concluding remarks.

2 A closer look at the patent system

Patents are monopoly-grants that hold for a time period of 20 years. During this period no one, apart from the patent holder, may freely make use of the technology embodied in the patent’s claims. Nevertheless, occasionally new ideas arrive, whose technological domain may well rest in a technological territory vaguely entrenched by the patent’s claims, in which case the issue of possible infringement arises, frequently accompanied by a counter claim of validity. This infringement differs from a direct copying and re-branding of one’s patented ideas, in as much as it progresses the prior art. The question of how and by how much the novel idea progresses prior art finds no equivalent in other forms of material-property. This is because, asserting property rights on one’s ideas is far from simple. An idea, contrasting land, can never be fully barricaded or entrenched. Therefore, the issue

³“According to U.S. patent law, the issuance of a patent does *no more than* confer a patent right that is “presumed” valid (35 U.S.C.A. Sec. 282) in that the final responsibility for validating or invalidating the patent resides with the courts.” (Choi, 1998, p.1249).

of infringement is largely a subjective one, resting on the decision of courts and it is up to the innovator, through litigation, to prove the merits of her innovation. In other words, patents convey imperfect property rights to technologies, rights that can only be asserted by courts. It has thus been argued that patents can be thought of as lotteries that carry some probability of being verified by a court; see Lemley and Shapiro (2005).

Two factors indisputably affect the enforcement probability of a patent. The first one is the scope (or breadth) of the patent, describing, through the claims that the patent office allows the patent holder to include, the patent's technological territory. The more extensive the claims are the more powerful a patent is, for it is harder to bypass it (e.g., by innovating around it) in developing a better and more advanced technology. Nonetheless, the claims of the patent themselves lack the ability to self-enforce. The power to enforce is vested in courts, which are the ultimate judges of the patents' merits. Therefore, the courts' attitude towards infringement is also a key determinant of patent validity.

The above factors are far from static. As far as the US is concerned, we have seen a drastic change in the past 20 years, both in court attitudes and patent scope. The catalyst for these changes was the Federal Courts Improvements Act that allowed for the formation of the CAFC. The CAFC raised the evidentiary standards required to challenge patent validity and broadened the interpretation of patent scope. In doing so the CAFC tilted the table towards patents making it easier to assert infringement. The data is revealing. Koenig (1980) provided validity and infringement data for district and circuit courts litigated patents during the years 1953-1978. He found that courts upheld the validity of patents in about 35% of the cases in which validity was an issue. This pattern was to drastically change with the introduction of the CAFC. Specifically, Harmon (1991) finds that from 1982 to 1990, the CAFC affirmed 90% of district court decisions holding patents to be valid and infringed, and reversed 28% of the judgments of invalidity and non-infringement. As a result, the overall probability that a litigated patent will be held to be valid has risen to 54%; see Allison and Lemley (1998).

This pro-patent stance is not cost-free. Merges (1999) points out that the expenses of a patent infringement court case can range from \$1 million to several millions, although Farrell and Merges (2004) show that as more money is at risk in the suit, litigation costs rise sharply. Taking a closer look at the total cost of patent litigation, Lerner (1994) reports that, from July 1989 to June 1990, 1318 patent related suits were initiated in the US Federal Court and approximately 3900 procedures within the US Patents and Trademarks Office (PTO). He estimates, based on historical costs, that these cases will

involve legal expenditures of about one billion 1991 dollars, which should be compared with expenditure on basic research of 3.7 billion by US firms in 1991.

It appears that litigation is an easier and less costly path to follow for firms with large patent portfolios. For example, Lanjouw and Schankerman (2004) find that having a larger portfolio of patents reduces the probability of filing a suit on any individual patent in the portfolio. As they note, “*for a (small) domestic unlisted company with a small portfolio of 100 patents, the average probability of litigating a given patent is 2%. For a company with a similar profile but with a moderate portfolio of 500 patents the figure drops to 0.5%. Thus, it is easier (less costly) to protect any given patent when that patent is part of a larger portfolio of patents*”. Furthermore, as they indicate, large firms (with large patent portfolios) have the experience and the ability to settle disputes by pooling or trading intellectual property. Therefore, if imperfect capital markets limit the capacity of smaller firms to finance litigation, larger firms may be able to extract better terms because they pose more credible litigation threats in confronting smaller firms.

On the other side of the Atlantic, even though the European Patent Office (EPO) grants patents using largely similar requirements to the PTO, it is stricter in granting a patent⁴ and it allows, through a post grant opposition mechanism, any interested party to challenge a patent at the EPO up until 9 months after the patent is granted. In fact, 8.2% of all EPO patents are challenged, and about one third are revoked; see Harhoff and Reitzig (2004). This procedure does not undermine the power of the member states’ courts, and if an opposing party, having lost its EPO opposition, wants to pursue its case at a national level it is free to do so, at a cost of anything between 50,000€ and 500,000€, depending on the country and the complexity of the case; see Ropski (1995). These differences limit the scope of EPO patents and reduce legal opposition.

3 A static benchmark

We consider two firms in the same industry, operating under a single line of patented, cumulative technology. Firm 1 is an established incumbent, holding an extensive patent portfolio. Firm 2, a potential entrant/startup, first decides whether to invest in an R&D that costs $C > 0$, if it decides against it the market stays unchanged and the game ends with payoffs of 0 to both firms. If firm 2 invests C , it develops a promising new

⁴As Graham and Harhoff (2006) suggest, even the most valuable US patents, those that are inviting costly litigation in US courts, are not being granted EPO patent protection in about 20% of cases.

technology and obtains one single-claim patent as a testimony to its innovativeness.⁵ Due to the cumulative nature of technology, however, firm 2's patent will be perceived as infringing on one or more of the incumbent's patents.

Firm 2 can generate a profit of V by commercializing the new technology. If firm 1 obtains access to the new technology and commercializes it in place of firm 2, it can generate a total profit of $V^* \geq V$ using its greater marketing experience. Firm 1 has three options described below.

The first option is to file a suit alleging that 2's technology is infringing on its patents, the outcome of which is uncertain. In line with Choi (1998), we model this as follows: Firm 1 wins the case with a commonly known probability $p \in (0, 1)$, in which case firm 1 reaps a profit of V^* by commercializing the new technology.⁶ With the complementary probability, $1 - p$, the court finds 2's technology non-infringing, in which case firm 2 commercializes it and reaps a profit of V . The payoff to the losing party is zero.

The second option is for firm 1 to seek a technology sharing agreement with 2 (i.e., a takeover deal or a licensing/cross-licensing agreement), which we model, in line with Crampes and Langinier (2002), as a Nash bargaining where the disagreement/threat points are the expected surpluses when firm 1 files an infringement lawsuit.

The third option is for firm 1 to do nothing, in which case firm 2's payoff is V from commercializing the technology and firm 1's payoff is 0. This option is clearly a dominated strategy, so we will not discuss it further. The structure of the game is common knowledge.⁷

The main exogenous variable of our model is p (and its determinants in the next section) where we interpret a higher p as reflecting a stronger stance of the court toward IP protection. The focus of analysis is on the extent to which p affects innovation incentives via its impact on the bargaining outcome.

In the static model described above, a strategy of firm 2 is whether to invest C or not, and that of firm 1 is whether to file a suit or to seek a Nash bargaining outcome. The subgame-perfect equilibrium of this game is obtained straightforwardly from the Nash

⁵Our results extend straightforwardly to the cases that investing C leads to an innovation with a known probability less than (rather than equal to) 1.

⁶The usual counter accusation of invalidity, if modeled here, would not change our results because it would entail no significant change in analysis other than reducing the value of filing a lawsuit for firm 1.

⁷As the foundations of an infringement suit must lay in the technical substitutability between patents, the option of a patent pool is not considered here because such substitutability would raise antitrust concerns as to the legality of this pool.

bargaining outcome as explained below. Specifically, the disagreement/threat points of firms 1 and 2 are, respectively, $d_1 = pV^*$ and $d_2 = (1-p)V$. Since V^* is the maximum value of the technology, the Nash bargaining set is defined as $B = \{(\bar{s}_1, \bar{s}_2) \in \mathfrak{R}_+^2 \mid \bar{s}_1 + \bar{s}_2 \leq V^*\}$ where \bar{s}_i denotes the bargaining share of firm $i = 1, 2$ (the bar above s_i is derogatory of the static framework). Given that B is compact and convex, there is a unique Nash bargaining outcome (\bar{s}_1, \bar{s}_2) that solves $\max_{(\bar{s}_1, \bar{s}_2) \in B} (\bar{s}_1 - d_1)(\bar{s}_2 - d_2)$, expressed as the following functions of p where $r = V/V^* \in [0, 1]$:

$$\bar{s}_1(p) = \frac{V^* + d_1 - d_2}{2} = \frac{1 + p(1+r) - r}{2} V^* \quad (1)$$

$$\bar{s}_2(p) = \frac{V^* - d_1 + d_2}{2} = \frac{1 - p(1+r) + r}{2} V^*. \quad (2)$$

Since $\bar{s}_1(p) > d_1$ and $\bar{s}_2(p) > d_2$ obviously hold, both firms will find it optimal to pursue a technology sharing agreement *a la* Nash bargaining, instead of litigation, once an innovation takes place. Anticipating such an agreement, a startup invests in R&D if and only if $\bar{s}_2(p) \geq C$.

The next proposition summarizes the findings in the static approach. Since the ownership of the patent *per se* does not alter the maximum value of the technology, V^* , the bargaining agreement need not take the form of a takeover, as it can equally well be attributed to licensing/cross-licensing.

Proposition 1: *In the static model, the Nash bargaining over a startup's innovation splits the total surplus V^* into $\bar{s}_1(p)$ for the incumbent and $\bar{s}_2(p)$ for the startup, as expressed in (1)-(2). Hence, a startup invests in R&D if and only if $\bar{s}_2(p) \geq C$. Stronger IP protection, i.e., a higher p , therefore, decreases (increases) the share of the startup (incumbent) via weakening (strengthening) its bargaining position, reducing the startup's innovation incentives.*

Note that some simplifying assumptions have been made for expositional ease, but these do not affect the main results as explained below. We do not consider preliminary injunctions which may enhance the plaintiff's negotiating power,⁸ however such an effect can be captured in our model via a greater p . We assume zero litigation cost, which is innocuous for our purpose because positive litigation costs would only strengthen our result by rendering out-of-court agreements more attractive. We also assume that justice is swift, which allows us to abstain from elaborating on the inessential details of the

⁸See Lemley and Shapiro (2007) for a formal analysis of this effect in a context where many patents read on a single product.

damages that the losing party needs to pay.⁹ Lastly, we assume that firm 1, if it wins the litigation, reaps a payoff of V^* , i.e., as if it had the sole right to commercialize the technology. In the previous version (2008) we assume that when 1 wins, some aspects of the infringing idea may find uses by potential rivals, lowering the actual value of the idea to firm 1 down to bV^* where $b \in [0, 1]$. All our results are valid for all values of $b \in [0, 1]$.

4 A dynamic approach

In this section we elaborate on the issues arising from the cumulative nature of technology by extending the model to infinite periods. In each period the static game described in the previous section is played as the stage game between a long-lived incumbent (firm 1) and a new potential startup (firm 2) that arrives at the market. To avoid the *replacement effect*, as in Bessen and Maskin (2007), we assume that V and V^* , the values that can be generated from commercializing the new technology, are incremental values. We stress here that, as will be illustrated in Section 5, our results are not driven by the disparity between V^* and V but rather by the additional bargaining power that expanding patent portfolios allow for.

If the incumbent acquires new patents through takeover deals, the technological territory covered by its patent portfolio expands and thus, as argued in the Introduction, the likelihood increases that it will prevail in patent-infringement suits. In this regard, we assume that legal power increases as the portfolio size gets bigger, but at a decreasing rate. That it increases at a decreasing rate is attested by Bessen and Meurer (2005) who observe decreasing returns to scale between the size of a software firm's patent portfolio and the probability of winning a patent litigation suit. Moreover, it is also a logical consequence of the fact that the chance of prevailing in court is bounded above by 1.

To capture this we re-define p , the probability of firm 1 winning an infringement suit, as a function of the degree of IP protection, denoted by $z \in (0, 1)$, and the size of 1's portfolio, measured by the number of patents in firm 1's portfolio. In particular, an increase in z (which can be considered as patent breadth) implies a tougher stance on infringement, increasing p .

⁹The yardstick used by courts in deriving damages is either the accumulated royalties resulting from a hypothetical licensing agreement (this is usually a per-period payment of 1-2% of the product's value), or the foregone profits from the sale of the infringing good. Both of these are minimal if justice is swift. Specifically, since a final product is yet to be developed there are no foregone profits, plus any foregone royalty payments cannot be central to the paper's argument because they have yet to accumulate.

To facilitate presentation, we make two indexing conventions. First, since the continuation game from any period is fully described by the size of 1's portfolio at the beginning of that period, with slight abuse of terminology we index the period by the size of 1's portfolio. Second, since what matters in the analysis is the accumulation of patents on top of the incumbent's initial portfolio, we index the size of the initial portfolio as the base size of 1, and each patent added to it increases the portfolio size by one. Hence, period 1 designates the initial period (of portfolio of size 1) and period $t > 1$ designates any period prior to which firm 1's portfolio size has reached t but no higher, i.e., firm 1 has added $t - 1$ patents to its initial portfolio. So long as firm 1 has added one patent every period from the initial period, our indexing coincides with the natural indexing of periods by natural numbers.¹⁰ Two consecutive periods are indexed the same, however, if the incumbent's portfolio did not grow in the first of the two periods. Thus, p is a function of z and t , which we denote as $p_z(t)$. As discussed earlier, we assume that

$$\partial p / \partial z > 0, \quad \partial p / \partial t > 0, \quad \text{and} \quad \partial^2 p / \partial t^2 < 0.$$

Our core argument starts with the observation that the benefits of a takeover venture beyond current bargaining for firm 1, as the added bargaining power (caused by the expansion of 1's portfolio) may mean better future deals. This suggests that the total surplus to bargain over can be larger than that in the static model. Consequently, firm 2 may rationally anticipate a larger bargaining share (compared to the static model), suggesting that dynamic incentives may induce innovations that would not have been possible in a static setting. This dynamic argument, unlike in the static model, implies that a takeover may be preferred to licensing because licensing (or cross-licensing) does not allow for the additional incremental value caused by the expansion in 1's patent portfolio.

The argument is not yet complete, because as takeover deals go on, i) firm 2's bargaining power weakens, and ii) the extra benefits of a takeover that accrues from future deals dwindles to zero due to decreasing returns to scale, rendering the total surplus/pie to bargain over to shrink to V^* , the size of the pie in the static model. Therefore, restricting, for example, the argument to innovations that would not materialize in the static setting (i.e. $C > \bar{s}_2(p)$), these two effects dictate that at some point firm 2 abstains from innovating as it would not recoup C . Such an end period upsets the potential equilibrium. To see this denote the last period that innovation is supposed to take place in equilibrium

¹⁰We do not model expiration of patents for expositional clarity. The effect of patent expiration is straightforward and can be seen easily when we have characterized the equilibrium, as briefly discussed at the end of this section.

as T . Then, the absence of future innovation reduces the bargaining surplus down to V^* in period T and consequently, firm 2's bargaining share would not cover C if it invested in period T . This implies no innovation in period T and, in addition, backward induction implies no innovation in all preceding periods, either.

This glitch is an artifact of the assumption that all startups have the same R&D cost C . We relax this assumption minimally by assuming that in each period there is a probability $\eta > 0$ that the arrived startup has an R&D cost smaller than C , normalized to 0 for expositional ease, and the realized value of R&D cost is private information but η is common knowledge. This is in line with Bessen and Maskin (2007) and amounts to assuming the existence of *Silicon Valley* startups that, contrasting their high-cost counterparts, innovate with minimum cost. As such firms will engage in R&D whenever they arrive, an innovation comes forth with at least probability η in every future period. Thus, in the last period T in which firm 2 would invest in R&D regardless of its R&D cost, the total surplus to bargain over is larger than V^* by at least a certain amount, and firm 2 may have an incentive to invest even when it is a high-cost firm.

To summarize, the order of moves in each period t is as follows. First, firm 2 arrives and decides whether to innovate or not contingent on its R&D cost which is 0 and C with probabilities η and $1 - \eta$, respectively. If 2 does not innovate, nothing happens until the next period starts. If 2 innovates, 1 decides whether to file a suit or pursue a technology-sharing agreement *a la* Nash bargaining. If a suit is filed, with probability $p_z(t)$ firm 1 wins and gets a surplus of V^* ; and with probability $1 - p_z(t)$ firm 2 wins and gets a surplus of V . The losing party has a surplus of 0. If an agreement is pursued, the Nash bargaining outcome results over the total surplus of V^* , plus, in case of a takeover, the additional benefits that would accrue to firm 1 in future deals due to its enlarged portfolio. We present our main analysis presuming that any technology-sharing agreement takes the form of a takeover,¹¹ then explain later how the results change when licensing agreements are allowed. The startup in each period maximizes the expected surplus of that period, net of innovation cost when relevant. The incumbent maximizes the expected present value of the stream of surpluses with a discount factor $\delta \in (0, 1)$.

Remark: Note that we do not explicitly model the possibility that the startup tries to build up its own portfolio via takeover deals with future startups. As long as we consider startups facing an incumbent with a large established portfolio, the value of pursuing this route would be low because the startup will continue to compete with the incumbent in the

¹¹Sale of patents to the incumbent will have the same effect as takeover.

product market as well as in future takeover deals, both of which will reduce the expected surplus. For example, modify the model so that in each period t any firm (**apart from the incumbent**) already in the market may file an infringement suit against the startup of period t or pursue a takeover deal. If the probability with which an existing firm with a single patent wins the infringement suit, \underline{p} , is sufficiently low relative to $p_z(1)$ (recall that $t = 1$ is our notational convention referring to the the patent portfolio size of the incumbent in the initial period), an existing firm with a single patent, having a very weak bargaining position, would only expect a small surplus from either pursuing a technology sharing agreement with the startup or from filing an infringement suit (against the startup or against the incumbent after a takeover deal is concluded between the incumbent and the startup). Thus, even after taking into account this option value, a startup would find a takeover deal preferable to the risky strategy of going through the litigation in the hope of building up its own portfolio. With V interpreted as including this option value, our model depicts market situations in which an incumbent has already accumulated a large patent portfolio so that the startups find it suboptimal, albeit feasible, to pursue building up its own portfolio by going through the litigation.¹²

We now provide a formal analysis of the dynamic model and characterize the (subgame-perfect) equilibrium.¹³ As we will show, there is a unique equilibrium and it exhibits the basic features elucidated above, namely, that the takeover deal *a la* Nash bargaining provides innovation incentives for high-cost startups, during an early stage of innovation dynamics at least. Such dynamic effects of inducing high-cost innovations would be best illustrated if a high-cost innovation would never be possible in a static situation. Hence, we first present our analysis in such environments, and then discuss other environments.

Thus, first we consider the case that $\bar{s}_2(p_z(1)) < C$, or equivalently,

$$\bar{s}_2(p_z(t)) < C \quad \text{for all } t \geq 1. \quad (3)$$

where $\bar{s}_2(\cdot)$ is firm 2's bargaining share in the static model as defined in (2). Let T denote, in an arbitrary equilibrium, the last period in which a high-cost startup innovates with a

¹²Admittedly, occasionally startups become dominant firms, however, these firms tend to rise to dominance in newly forming markets (e.g., Google, Genetech), rather than “win” it over from established incumbents.

¹³Firm 1 decides to litigate or bargain without knowing the R&D cost of firm 2, so technically speaking the subgame-perfectness does not require rationality of the incumbent's decision. However, the R&D cost of the startup is sunk at this point, so it does not affect the continuation game. Hence, in the spirit of subgame-perfectness, we require that each choice of the incumbent be optimal in the continuation game.

positive probability, allowing for the possibility that $T = 0$, i.e., a high-cost startup never innovates. T is our point of departure in the analysis, and for notational purposes, in the sequel a hat on top of a variable is derogatory of all $t \leq T$ periods, and absence of a hat denotes all $t > T$ periods. Given that $T < \infty$ exists (indicating that in equilibrium a high-cost startup would not innovate indefinitely) as is proved in Proposition 3 below, for $t \geq T + 1$, let $X(t)$ denote the value of firm 1 at the beginning of period t . Then,

$$X(t) = (1 - \eta)\delta X(t) + \eta(s_1(t) + \delta X(t + 1)) \quad (4)$$

because, a) if a high-cost firm arrives with probability $1 - \eta$, there is no innovation and firm 1's value in the next period is the same as that in the current period (i.e. $X(t)$) and, b) if a low-cost firm arrives with probability η , firm 1 captures the bargaining surplus over the current innovation, $s_1(t)$, plus its value in the next period which is $X(t + 1)$.

Focusing on equation (4), for $t > T$ the total surplus that a startup's innovation generates is maximized when firm 1 commercializes it, adding it to its portfolio. The total surplus it brings forth in this case is $V^* + \delta(X(t + 1) - X(t))$, which is the size of the pie on the bargaining table. If the case is litigated, since both parties must accept the court's decision, there is no takeover deal. Therefore, the threat points are the court outcomes, $d_1 = p_z(t)V^*$ and $d_2 = (1 - p_z(t))V$. Since the Nash bargaining set in this case is $B(t) = \{(s_1, s_2) \in \mathfrak{R}_+^2 \mid s_1 + s_2 \leq V^* + \delta(X(t + 1) - X(t))\}$, the Nash bargaining outcome (s_1, s_2) that solves $\max_{(s_1, s_2) \in B(t)} (s_1 - d_1)(s_2 - d_2)$ is calculated as,

$$s_1(t) = \frac{1 + p_z(t)(1 + r) - r}{2} V^* + \frac{\delta(X(t + 1) - X(t))}{2} \quad (5)$$

$$s_2(t) = \frac{1 - p_z(t)(1 + r) + r}{2} V^* + \frac{\delta(X(t + 1) - X(t))}{2} \quad (6)$$

where $r = V/V^* \in [0, 1]$. Plugging $s_1(t)$ back into equation (4) and rearranging, we get

$$X(t + 1) - X(t) = \frac{2(1 - \delta)}{3\delta\eta} X(t) - \frac{1 + p_z(t)(1 + r) - r}{3\delta} V^*, \quad (7)$$

a difference equation that characterizes the sequence $X(t)$ for $t > T$. Since the value of additional patent diminishes to 0 as $t \rightarrow \infty$, it turns out that this sequence increases and converges, as formalized in the next result. Although $X(t)$ is pertinent for $t > T$, it proves useful to treat it as a function defined for all natural numbers $t \geq 1$.

Proposition 2: *The sequence $\{X(t)\}$ defined by (7) is unique, monotonically increases at a decreasing rate, i.e., $X(t) - X(t - 1) > X(t + 1) - X(t) > 0$ for all $t > 1$, and converges to*

$$X(\infty) = \frac{1 - r + p_z(\infty)(1 + r)}{2(1 - \delta)} V^* \eta \quad \text{as } t \rightarrow \infty. \quad (8)$$

Proof: See Appendix.

Proposition 3: *If (3) holds, in any equilibrium there exists an earliest period $T < \infty$ such that a high-cost startup does not innovate in any period $t > T$.*

Proof: See Appendix.

Let $\hat{X}(t)$ denote firm 1's value at the beginning of period t for $t \leq T$. Presuming that a high-cost startup innovates for sure in period T , firm 1's value at the beginning of T is

$$\hat{X}(T) = \hat{s}_1(T) + \delta X(T+1) \quad (9)$$

where $\hat{s}_1(T)$ denotes the bargaining share that it derives over the current innovation. Given that the total surplus to bargain over is $V^* + \delta(X(T+1) - \hat{X}(T))$ and the threat points are $d_1 = p_z(T)V^*$ and $d_2 = (1 - p_z(T))V$ in period T , we calculate the Nash bargaining outcome $(\hat{s}_1(T), \hat{s}_2(T))$ as

$$\hat{s}_1(T) = \frac{1 + p_z(T)(1+r) - r}{2} V^* + \frac{\delta(X(T+1) - \hat{X}(T))}{2} \quad \text{and} \quad (10)$$

$$\hat{s}_2(T) = \frac{1 - p_z(T)(1+r) + r}{2} V^* + \frac{\delta(X(T+1) - \hat{X}(T))}{2}. \quad (11)$$

Plugging $\hat{s}_1(T)$ into equation (9), we derive $\hat{X}(T)$ in terms of $X(T+1)$ as

$$\hat{X}(T) = \left(1 + \frac{\delta}{2}\right)^{-1} \cdot \left(\frac{3\delta}{2}X(T+1) + \frac{1 + p_z(T)(1+r) - r}{2}V^*\right). \quad (12)$$

Furthermore, rearranging equation (7) for $t = T$ we get

$$X(T) = \left(1 - \delta + \frac{3\delta\eta}{2}\right)^{-1} \cdot \left(\frac{3\delta}{2}X(T+1) + \frac{1 + p_z(T)(1+r) - r}{2}V^*\right)\eta. \quad (13)$$

Since $1 - \delta + \frac{3\delta\eta}{2} > (1 - \delta)\eta + \frac{3\delta\eta}{2} = (1 + \frac{\delta}{2})\eta$, it follows from (12) and (13) that $\hat{X}(T) > X(T)$, hence $\hat{s}_2(T) < s_2(T)$. Note that $s_2(T) > C$ must hold for otherwise a high-cost startup would not innovate in period T because $\hat{s}_2(T) < s_2(T)$ would imply $\hat{s}_2(T) < C$. Also, $s_2(T+1) \leq C$ must hold for otherwise a high-cost startup would innovate in period $T+1$. Since $s_2(t)$ monotonically decreases in t from (6) and Proposition 2, therefore, we deduce that

$$T = \max\{t \mid s_2(t) > C\}. \quad (14)$$

If $\hat{s}_2(T) \geq C$, then a high-cost startup would innovate in period T as presumed. But, it is also possible that $\hat{s}_2(T) < C < s_2(T)$, in which case a high-cost startup would not innovate in period T . This problem is resolved when mixed strategies are considered: if a

high-cost startup invests with an appropriate probability, $\hat{X}(T)$ gets reduced, pushing up $\hat{s}_2(T)$ to a level equal to C so that the startup is indifferent between investing and not. Specifically, if $\hat{s}_2(T) < C < s_2(T)$ we redefine $\hat{X}(T)$ and $\hat{s}_1(T)$ as $\hat{X}(T, a)$ and $\hat{s}_1(T, a)$ that solve

$$\hat{X}(T, a) = (\eta + a)(\hat{s}_1(T, a) + \delta X(T + 1)) + (1 - \eta - a)\delta \hat{X}(T, a), \quad (15)$$

$$\hat{s}_1(T, a) = \frac{1 + p_z(T)(1 + r) - r}{2} V^* + \frac{\delta(X(T + 1) - \hat{X}(T, a))}{2}, \text{ and} \quad (16)$$

$$\hat{s}_2(T, a) = \frac{1 - p_z(T)(1 + r) - r}{2} V^* + \frac{\delta(X(T + 1) - \hat{X}(T, a))}{2}, \quad (17)$$

for some $a \in (0, 1 - \eta)$ so that, in particular,

$$\hat{X}(T, a) = \left(1 - \delta + \frac{3\delta(\eta + a)}{2}\right)^{-1} \cdot \left(\frac{3\delta}{2} X(T + 1) + \frac{1 + p_z(T)(1 + r) - r}{2} V^*\right)(\eta + a). \quad (18)$$

As a increases from 0 to $1 - \eta$, $\hat{X}(T, a)$ increases from $X(T)$ of equation (13) to $\hat{X}(T)$ of equation (12). Analogously, $\hat{s}_2(T, a)$ decreases from $s_2(T)$ to $\hat{s}_2(T)$. Since $\hat{s}_2(T) < C < s_2(T)$, it follows that there exists a unique value of $a \in (0, 1 - \eta)$, denoted by $\hat{a}(T)$, such that $\hat{s}_2(T, \hat{a}(T)) = C$. Thus, if a high-cost startup were to invest C with probability $\frac{\hat{a}(T)}{1 - \eta}$ in period T , its bargaining share would be $\hat{s}_2(T, \hat{a}(T)) = C$, ensuring that a high-cost startup is indifferent between innovating and not and thus, justifying the mixed strategy.

We have specified above the unique equilibrium behavior in period T , according to which a high-cost startup innovates with a positive probability. This, however, does not warrant that an innovation takes place for sure in all preceding periods $t < T$ because, as before, the prospect of sure innovation in period t may increase the value $\hat{X}(t)$ too much and thereby, reduce the marginal value of an additional patent ($\hat{X}(t + 1) - \hat{X}(t)$) too low a level to generate large enough a bargaining share for a high-cost startup to recoup its R&D cost C . In such periods, by recursively applying the same logic as above to periods $t = T - 1, T - 2$, and so on, we obtain a unique equilibrium strategy in which a high-cost firm innovates with a probability strictly between 0 and 1. Since this process is analogous to finding the probability $\hat{a}(T)$ explained above, we refer the details to a previous version of this paper (2008). Consequently, in the unique equilibrium high-cost startups innovate with a strictly positive probability in all periods $t \leq T$, but never innovate in later periods. This completes characterization of the unique equilibrium for the cases that satisfy (3).

We now discuss the alternative case that $\bar{s}_2(p_z(1)) \geq C$. There are two subcases to consider, namely, $\bar{s}_2(p_z(\infty)) < C$ and $\bar{s}_2(p_z(\infty)) \geq C$. When $\bar{s}_2(p_z(\infty)) < C$ it is straightforward to see that there exists a unique equilibrium analogous to the one

characterized above. Specifically, high-cost startups innovate for sure in all periods t such that $\bar{s}_2(p_z(t)) \geq C$, because $\hat{X}(t)$ increases in t and consequently,

$$\hat{s}_2(t) = \bar{s}_2(p_z(t)) + \frac{\delta(\hat{X}(t+1) - \hat{X}(t))}{2} > C. \quad (19)$$

Then, since the increase in $\hat{X}(t)$ slows down and, for some t , $\bar{s}_2(p_z(t))$ will eventually dip below C , high-cost startups stop innovating from a certain period. If $\bar{s}_2(p_z(\infty)) \geq C$, on the other hand, a high-cost startup innovates in every period in the unique equilibrium of the dynamic model because $\hat{s}_2(t) > \bar{s}_2(p_z(t))$ as per (19) and $\bar{s}_2(p_z(t)) > \bar{s}_2(p_z(\infty))$. Now we characterize the unique equilibrium of the dynamic model in the next theorem.

Theorem 4: *The dynamic model has a unique equilibrium. If $\bar{s}_2(p_z(\infty)) \geq C$, in this equilibrium a startup innovates for sure in every period regardless of its R&D cost; If $\bar{s}_2(p_z(\infty)) < C$, on the other hand, there is a critical period $T < \infty$ such that a high-cost startup innovates with a positive probability in every period $t \leq T$ but not in periods $t > T$, while a low-cost startup innovates for sure in every period. In either case, when there is an innovation the incumbent reaches a takeover deal with the innovator.*

An interesting policy-relevant question is what is the optimal level of IP protection, z , that provides the innovation incentives for startups for longest. Although an algebraic answer is hard to obtain due to the recursive nature of the solution and the discontinuity at $t = T$, our simulation results (summarized in Section 5) confirm the following intuition: If z is excessive, the marginal protective power that an extra patent brings to the incumbent is large initially but quickly dwindles as a result of accumulating its power too rapidly, killing off the positive effect on startup innovation prematurely. If z is flimsy, on the other hand, the marginal protective power of an extra patent can be too small and its impact on the startup's innovation incentives limited. Consequently, the optimal level of IP protection tends to be at a moderate level.

We have carried out our analysis presuming that any technology-sharing agreement is restricted to a takeover deal, i.e., licensing was not considered. Since licensing (lacking the added advantages accruing to firm 1 from future dealings) fails to increase the innovation's total value beyond V^* , one can see that the Nash bargaining outcome of a licensing deal is the same as the static model's bargaining outcome, $\bar{s}_1(p_z(t))$ and $\bar{s}_2(p_z(t))$. Thus, if $\bar{s}_2(p_z(1)) < C$ licensing in any period t would not cover C for the startup, allowing only low-cost innovations, while if $\bar{s}_2(p_z(\infty)) \geq C$ it would allow high-cost innovations in every period t . In either case, reaching a takeover deal instead of licensing would not affect the

startups' innovation decisions. However, with takeover deals firm 1 anticipates a larger surplus due to its enhanced future bargaining positions and, furthermore, a part of this extra surplus accrues to the current startup firm (at the expense of the startups in future deals). Therefore, the equilibrium outcome does not change even if licensing is allowed.

If $\bar{s}_2(p_z(1)) \geq C > \bar{s}_2(p_z(\infty))$, on the other hand, there is room for licensing. For example, after expanding its portfolio via takeovers to the largest size, say \bar{t} , such that $\bar{s}_2(p_z(\bar{t})) \geq C$, the incumbent may obtain access to new technology through licensing in all subsequent periods, so that startups innovate forever regardless of their R&D cost. Relative to when licensing is disallowed, this would bring about more innovations but the bargaining share of the incumbent would be smaller for the periods in which licensing deals will be reached: This is because \bar{t} is smaller than T , the largest portfolio size consistent with high-cost innovation as explained above. Consequently, the incumbent may be able to do better by expanding its portfolio size to T via takeovers, then selectively releasing some of the patents in its portfolio to avoid discouraging startup innovations by becoming too powerful a potential plaintiff. This practise is reminiscent of the recent trend of patent donations: in the last few years firms such as DuPont, Lubrizol, Eastman Chemicals, and General Motors have given away patents with an estimated value of hundreds of millions of dollars. An interesting alternative interpretation of T , therefore, is the optimal patent length that, by constraining the incumbent's portfolio size below a threshold, would allow for the arrival of high-cost innovations *ad infinitum*.

5 Simulation and comparative statics

In light of the diminishing marginal protective power of patents underlying our theoretical results, prior to the simulation we need to address how $p_z(t)$ changes with t and z . Empirical estimates of the marginal protective ability of patents are scarce and inconclusive. Lanjoun and Schankerman (2004), who look at how patent portfolios help reduce a firm's probability of facing litigation, are one of the few that examine how patent portfolios affect litigation. They find the marginal protective power to be positive but slowing down. We capture this through $p_z(t) = 1 - (1 - z)^t$. To provide an example (in line with the magnitudes of z we find), when $z = .007$ a firm with a portfolio made up of 100 patents stands a 50% chance of winning its case, and an increase of 1 patent raises this by .34%.

Normalizing $V = 1$, we argue our case for $r = 1$, and $C = 1.0001$. An $r = 1$ allows for results that are not driven by the disparity between V and V^* , and, since

$\bar{s}_2(p) \leq \bar{s}_2(0) = 1$, by choosing $C > 1$ we ensure that a) innovation by high-cost startups may only be possible in a dynamic model, and b) that IP protection is necessary for such innovation because $p_z(t)$ is constant at 0 for all t if $z = 0$, erasing any dynamic effect. We need to mention that $r = 1$ and $C > V$ were chosen to demonstrate these points, and should not be construed as representative of real situations. For more relaxed parameter values, high-cost innovation will be sustained more broadly than our simulation reported below. In addition, high-cost innovation is not socially desirable for the chosen parameter values because $C > V^*$, but it will be if $r < 1.0001^{-1}$. In terms of δ and η , we initially simulate the model for $\delta = .97$ and $\eta = .2$ and then for $\delta = .98$ and $\eta = .8$.

Concentrating on the $t > T$ periods, for given z , we simulate the unique sequence $\{X(t)\}$ defined by (7), which converges to (8). From this sequence, through equation (6), is derived a convergent sequence $s_2(t)$, as illustrated in Figure 3 for $z = .007$. Then, the period T is obtained by identifying the last period for which $s_2(t) > C$.

To examine how z affects T , we find the values of T for z 's between .0001 and .01 in 20 steps of .0005. Figure 4 shows how T changes as z increases when $\delta = .97$ and $\eta = .2$ (the lower graph) and when $\delta = .98$ and $\eta = .8$ (the upper graph).

Both graphs are quasiconcave, in particular, T initially increases with z , then decreases as z increases further. Specifically, if z is high then $X(t)$ converges quickly, driving the future benefits from an extra takeover to nil and thus, halting the positive effect on startup innovation prematurely. For small z 's, on the other hand, an extra patent increases $p_z(t)$ and $X(t)$ only marginally, failing to sufficiently increase $s_2(t)$ as to allow for a high T . Needless to say, the precise relationship between z and T changes as other details of the specification change. However, the quasiconcavity with an interior peak prevails in all our simulations so long as $\bar{s}_2(0) < C$ and high-cost innovation is possible at all, as partly demonstrated in Figure 4 explained above. Furthermore, in our analysis both firms are assumed to have equal bargaining power, however, endowing different bargaining power would lead to different levels of T . Lowering C would also lead to a higher T .

Turning our attention to $t \leq T$, for $\delta = .97$, $\eta = .2$ (and for the same z 's used above) we employ equations (11)-(12) to derive $\hat{X}(T)$ and $\hat{s}_2(T)$. If $\hat{s}_2(T)$ is less than C , using mixed strategies, we find $\hat{a}(T)$ by equating (17) to C , and derive $\hat{X}(T, \hat{a}(T))$ from equation (18). Using the same routine we derive the equilibrium value of $\hat{a}(t)$ for each $t \leq T$, as reported in Figure 5 for each of the 20 different values of z . Due to the discreteness, a range of z provides the highest T as shown in Figure 5. Among these values of z , those with higher values of $\hat{a}(t)$ induce more startup innovation on average in each period $t \leq T$.

In Figure 5, the middle values in the range of z for the highest T tend to have higher values of $\hat{a}(t)$ although no single value of z is pinned down as having the highest $\hat{a}(t)$ for all $t \leq T$.

Finally, we report some comparative statics results by examining how η , and δ affect the graphs in Figure 4. Starting with δ , Figure 6 plots the graphs as we vary δ from .9 to .99 in 10 steps (changing z as before), which indicate that T increases as δ increases. Lastly, Figure 7 plots the graphs as η varies from .1 to 1 in 10 steps, which indicate that T increases (at a decreasing rate) as η increases.

6 Conclusions

In the 1990's a vibrant literature analyzed the rate of IP protection (in terms of patent breadth *vs.* length) that minimizes the effects of the monopoly that patents imply while offering sufficient R&D incentives. In this paper we depart from this tradition focusing instead on the dynamic effects of IP protection on startup innovation. We argue that positive but not excessive IP protection may foster takeover agreements between the incumbent and startup innovator, the prospect of which motivates the startup's entrepreneurial activity in the first place. Since the benefits of a takeover venture beyond the current invention via strengthened bargaining position in future takeover deals due to an enlarge patent portfolio, innovation activities may be less active when only licensing agreements are allowed or when there is no IP protection. On the other hand, excessive IP protection accumulates the incumbent's bargaining power too quickly and kills off the startup's innovation activities prematurely. We demonstrate this intuition by simulation results that exhibit an inverse U-shape relationship between the number of sustainable innovations and the level of IP protection.

Our theory can help explain the increase in takeovers we have witnessed since the 1980's, an increase that coincides with a major shift in US patent policy, namely, the formation of a single patent's court, in place of many appellate courts with diverging attitudes towards infringement. Such a legal apparatus is absent from the EU, where questions of enforcement, validity and revocation are dealt with by national courts that have varying attitudes towards infringement. The EU Commission has been advocating the creation of a central patent dispute court, in the hope of simplifying the IP framework in the EU. Our analysis supports this initiative suggesting that, provided the court keeps a balanced approach, it can spur innovation activities by startup entrepreneurs.

We have labored under the assumption that there is only one firm interested in the startup. The struggle between Microsoft and Google for a share of Facebook suggests that this is not always the case. How the bargaining will get resolved between a startup and competing incumbents is unclear, because (setting aside competition law issues) there is a variety of different strategies that the competitors may follow, which venture beyond the scope of this paper. Subsequently, we leave such a question as a further agenda.

Appendix

Proof of Proposition 2: First, note that $X(t)$ is bounded below (by 0) and above because maximum surplus in each period is bounded and $\delta < 1$. If $X(t+1) \leq X(t)$, then the right hand side of equation (7) would be non-positive and, furthermore, its value would strictly decrease when evaluated for $t+1$ because $X(t+1) \leq X(t)$ and $p_z(t+1) > p_z(t)$. This would mean that $X(t+2) - X(t+1) < X(t+1) - X(t) \leq 0$. Applying the same argument repeatedly, we deduce that if $X(t+1) \leq X(t)$ then the sequence should decrease forever at an increasing rate after t , which is a contradiction because the sequence is bounded below. Hence, we conclude that $X(t+1) - X(t) > 0$ for all t . Since the sequence is bounded above, it further follows that it must converge. The limit value, $X(\infty)$ in (8), is obtained by setting $X(t+1) = X(t)$ and $p_z(t) = p_z(\infty)$ in equation (7) and solving for $X(t)$.

To show uniqueness, suppose to the contrary that there are two sequences, $\{X(t)\}$ and $\{X'(t)\}$, that satisfy (7), such that $X'(t') = X(t') + \gamma$ for some $\gamma > 0$ and t' . By (7), we have $X'(t'+1) = X(t'+1) + (1 + \frac{2(1-\delta)}{3\delta\eta})\gamma > X(t'+1) + \gamma$ and by repeating the same calculation, $X'(t) > X(t) + \gamma$ for all $t \geq t'$. This is impossible because both sequences should converge to the same limit as proved above, proving the uniqueness.

Finally, to show that $X(t) - X(t-1) > X(t+1) - X(t)$, note from equation (7) that

$$X(t+1) - X(t) - (X(t) - X(t-1)) = \frac{2(1-\delta)}{3\delta\eta}(X(t) - X(t-1)) - \frac{(p_z(t) - p_z(t-1))(1+r)}{3\delta}V^*. \quad (20)$$

If $X(t+1) - X(t) \geq X(t) - X(t-1)$ for some t , it would follow from equation (20) that $X(t+2) - X(t+1) \geq X(t+1) - X(t)$ because $0 < p_z(t+1) - p_z(t) < p_z(t) - p_z(t-1)$ due to the assumption that $\partial^2 p / \partial t^2 < 0$. Furthermore, $X(t+1) - X(t)$ would increase in t by repeated application of the same argument. This is impossible because the sequence $X(t)$ converges as shown above, hence we conclude that $X(t) - X(t-1) > X(t+1) - X(t)$. *Q.E.D.*

Proof of Proposition 3: To reach a contradiction, suppose to the contrary that in an equilibrium there is an arbitrarily large t such that a high-cost startup innovates with a positive probability in period t . Note that a takeover deal will be reached if an innovation takes place in period t , for otherwise a high-cost startup would not innovate because $\bar{s}_2(p_z(t)) < C$ by (3). Let \hat{X}_t denote firm 1's value at the beginning of period t , and let α_t denote the probability that an innovation takes place in period t . Since a low-cost startup always innovates, $\alpha_t = \eta + (1 - \eta)a_t \geq \eta$ where a_t is the probability that a high-cost startup innovates in period t . Then, $\hat{X}_t = \alpha_t(\hat{s}_{1t} + \delta\hat{X}_{t+1}) + (1 - \alpha_t)\delta\hat{X}_t$ where $\hat{s}_{1t} = \frac{1+p_z(t)(1+r)-r}{2}V^* + \frac{\delta(\hat{X}_{t+1}-\hat{X}_t)}{2}$, so that

$$\hat{X}_t = \frac{\alpha_t}{1 - \delta + \frac{3}{2}\delta\alpha_t} \left(\frac{3}{2}\delta\hat{X}_{t+1} + \bar{s}_1(p_z(t)) \right). \quad (21)$$

If $\hat{X}_t \geq \hat{X}_{t+1}$, firm 2's bargaining share in period t , $\hat{s}_{2t} = \bar{s}_2(p_z(t)) + \frac{\delta(\hat{X}_{t+1}-\hat{X}_t)}{2}$, is less than C because $\bar{s}_2(p_z(t)) < C$ by (3), and consequently, $\alpha_t = \eta$. Since $\alpha_{t+1} \geq \eta$ and $p_z(t+1) \geq p_z(t)$, therefore, $\hat{X}_t \geq \hat{X}_{t+1} < \hat{X}_{t+2}$ would imply

$$\frac{\alpha_t}{1 - \delta + \frac{3}{2}\delta\alpha_t} \left(\frac{3}{2}\delta\hat{X}_{t+1} + \bar{s}_1(p_z(t)) \right) < \frac{\alpha_{t+1}}{1 - \delta + \frac{3}{2}\delta\alpha_{t+1}} \left(\frac{3}{2}\delta\hat{X}_{t+2} + \bar{s}_1(p_z(t+1)) \right),$$

contradicting the presumption that $\hat{X}_t \geq \hat{X}_{t+1}$ according to (21). Hence, we deduce that if $\hat{X}_t \geq \hat{X}_{t+1}$ then $\alpha_t = \eta$ and $\hat{X}_{t+1} \geq \hat{X}_{t+2}$, and by repeatedly applying the same logic, $\alpha_{t'} = \eta$ for all $t' > t$. Since this contradicts to the supposed equilibrium, we conclude that $\hat{X}_t < \hat{X}_{t+1}$ for all t . Since firm 1's value is bounded above, it further follows that $\hat{X}_{t+1} - \hat{X}_t \rightarrow 0$ as $t \rightarrow \infty$, which in turn implies that $\hat{s}_{2t} \rightarrow \bar{s}_2(p_z(\infty))$ as $t \rightarrow \infty$, contradicting the presumption that a high-cost startup innovates with a positive probability indefinitely. Q.E.D.

References

- Allison, J., and Lemley, M., (1998), 'Empirical Evidence on the Validity of Litigated Patents', *AIPLA Quarterly Journal*, **26**, pp. 185-275.
- Aoki, R., and Hu, J., (1998), 'Licensing v. Litigation: Effects of the Legal System on the Incentives to Innovate', *Journal of Economics and Management Science*, **8**, pp. 133-160.
- Bessen, J., and Meurer, M., (2005), 'The Patent Litigation Explosion', Boston University, School of Law, Working Paper No. 05-18.
- Bessen, J., and Maskin, E., (2007), 'Sequential Innovation, Patents, and Imitation', *The RAND Journal of Economics*, forthcoming.
- Boldrin, M., and Levine, D., (2002), 'The Case against Intellectual Property', *American Economic Review*, **92**, May, pp. 209-212.

- Chang, H., (1995), ‘Patent Scope, Antitrust Policy, and Cumulative Innovation’, *RAND Journal of Economics*, **26**, pp. 34-57.
- Chesbrough, H. (2003), *Open Innovation: The New Imperative for Creating and Profiting from Technology*, HBS Press Book, USA.
- Choi, J., P., (1998), ‘Patent Litigation as an Information-Transmission Mechanism’, *American Economic Review*, **88**, pp 1249-1263.
- Crampes, C., and Langinier, C., (2002), ‘Litigation and Settlement in Patent Infringement Cases’, *RAND Journal of Economics*, **33**, pp. 258-274.
- Dessyllas, P., and Hughes, A., (2005), ‘R&D and Patenting Activity and the Propensity to Acquire in High Technology Industries’, ESRC Centre for Business Research, University of Cambridge Working Paper No. 298.
- Farrell, J., and Merges, R., (2004), ‘Incentives to Challenge and Defend Patents: Why Litigation Won’t Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help’, *Berkeley Technology Law Journal, Annual Review of Law and Technology*, **19**, No. 1.
- Farrell, J., and Shapiro, C., (2006), ‘How Strong are Weak Patents?’, University of California, Working Paper.
- Graham, S. and Harhoff, D., (2006), ‘Can Post-Grant Reviews Improve Patent System Design? A Twin Study of US and European Patents’, CEPR DP 5680.
- Green, J., and Scotchmer, S., (1995), ‘On the Division of Profit in Sequential Innovation’, *RAND Journal of Economics*, **26**, pp. 20-33.
- Harhoff, D., and Reitzig, M., (2004), ‘Determinants of Opposition Against EPO Patent Grants—The Case of Biotechnology and Pharmaceuticals’, *International Journal of Industrial Organization*, **22**, pp. 443-480.
- Harmon, R., (1991), ‘Patents and the Federal Circuit’, Bureau of National Affairs, Washington, D.C.
- Hopenhayn, H., Llobet, G., and Mitchell, M., (2006), ‘Rewarding Sequential Innovators: Prizes, Patents, and Buyouts’, *Journal of Political Economy*, **114**, pp. 1041-68.
- Koenig, G., (1980), *Patent Invalidity: A Statistical and Substantive Analysis*, Clark Boardman, New York, USA.
- Lanjouw, J., and Schankerman, M., (2004), ‘Protecting Intellectual property Rights: Are Small Firms Handicapped?’, *Journal of Law and Economics*, **47**, pp. 45-74.
- Lemley, M., and Shapiro, C., (2005), ‘Probabilistic Patents’, *Journal of Economic Perspectives*, **19**, pp. 75-98.
- Lemley, M., and Shapiro, C., (2007), ‘Patent Holdup and Royalty Stacking’, *Texas Law Review*,

85, pp. 1991-2049.

Lerner, J., (1994), 'The Importance of Patent Scope: An Empirical Analysis', *RAND Journal of Economics*, **25**, pp. 319-333.

Lerner, J., and Tirole, J., (2004), 'Efficient Patent Pools', *American Economic Review*, **94**, pp 691-711.

Llobet, G., (2003), 'Patent Litigation When Innovation is Cumulative', *International Journal of Industrial Organization*, **21**, pp. 1135-1157.

Merges, R., (1999), 'As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform', *Berkeley Technology Law Journal*, **14**, pp. 577-615.

Meurer, M., (1989), 'The Settlement of Patent Litigation.' *RAND Journal of Economics*, **20**, pp. 77-91.

O'Donoghue, T., Scotchmer, S., and Thisse, J., (1998), 'Patent Breadth, Patent Life, and the Pace of Technological Progress', *Journal of Economics & Management Strategy*, **7**, pp. 1-32.

Panagopoulos, A., and Park, I.-U., (2008), 'Patent Protection, Takeovers, and Startup Innovation: A Dynamic Approach', CMPO WP 08/201, University of Bristol.

Scotchmer, S., (1996), 'Protecting Early Innovators: Should Second Generation Products be Patentable?', *RAND Journal of Economics*, **27**, pp. 322-331.

Ropski, G. (1995), *Enforcing a Global Patent Portfolio*, Butterworths, London.

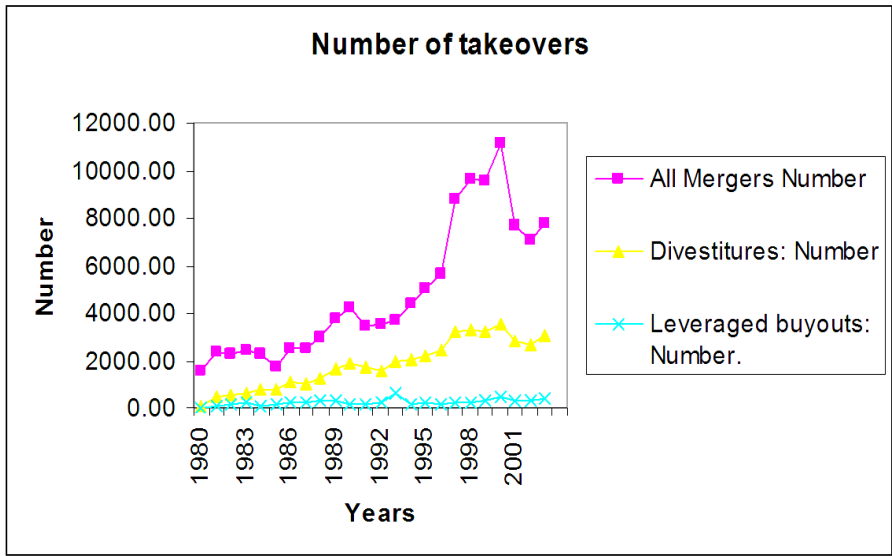


Figure 1: Number of Mergers, Divestitures, and Leveraged buyouts from 1980 to 2003. Source: Statistical Abstracts of the United States.

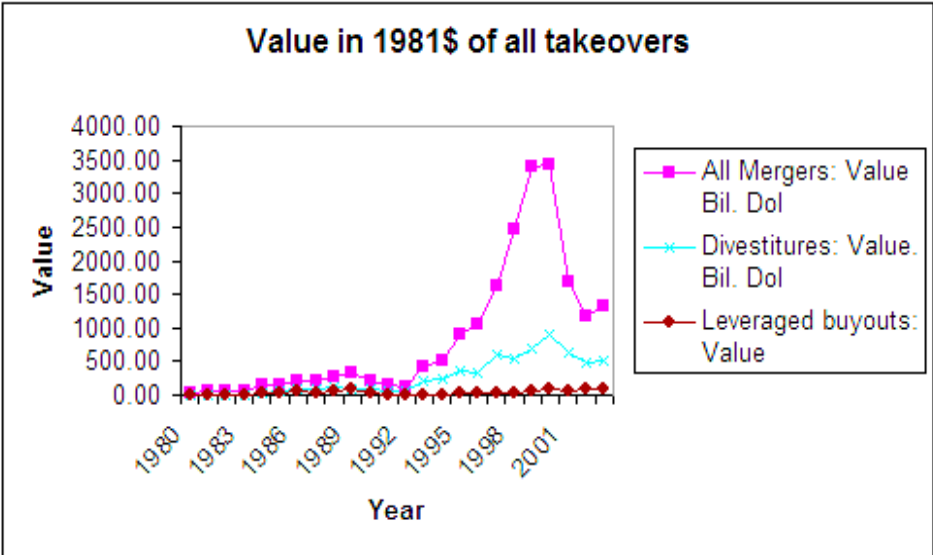


Figure 2: Value of Mergers, Divestitures, and Leveraged buyouts from 1980 to 2005. Source: Statistical Abstracts of the United States.

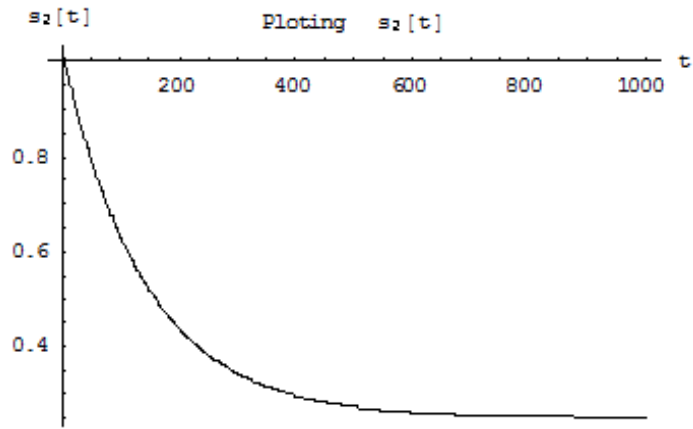


Figure 3: The simulated $s_2(t)$ for $z = .007$, $\delta = .97$ and $\eta = .2$

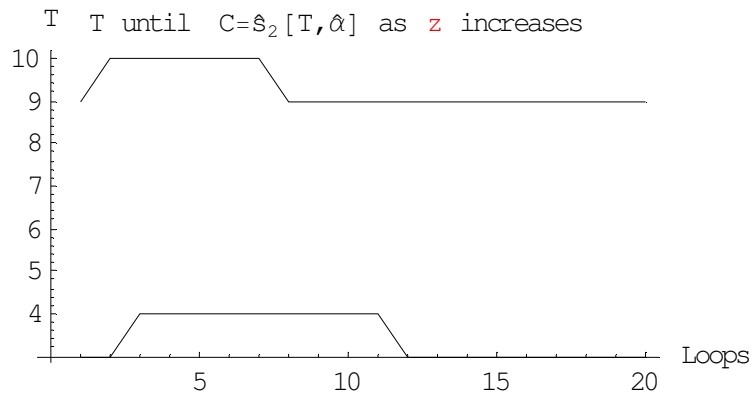


Figure 4: The simulated effect of z on T , for z 's between .0001 and .01 in 20 steps of .0005. The lower graph plots the T 's for $\delta = .97$ and $\eta = .2$, while the upper graph does the same for $\delta = .98$ and $\eta = .8$

The alphas					
z	T	t=1	t=2	t=3	t=4
0.0005	3	0.44504	0.100914	0.00589	
0.001	3		0.305085	0.077427	
0.0015	4		0.334985	0.08551	0.003332
0.002	4		0.345239	0.087752	0.004101
0.0025	4		0.349353	0.088245	0.004061
0.003	4		0.350775	0.08799	0.003679
0.0035	4		0.350787	0.087351	0.003123
0.004	4		0.349975	0.08649	0.002468
0.0045	4		0.348642	0.085493	0.001753
0.005	4		0.346964	0.084408	0.001
0.0055	4		0.345049	0.083263	0.000222
0.006	3		0.343123	0.082158	
0.0065	3		0.341139	0.081057	
0.007	3		0.339069	0.079939	
0.0075	3		0.336934	0.078808	
0.008	3		0.334751	0.07767	
0.0085	3		0.332533	0.076527	
0.009	3		0.33029	0.075383	
0.0095	3		0.328029	0.074238	
0.01	3		0.325755	0.073095	

Figure 5: The equilibrium value of $\hat{a}(t)$ for each $t \leq T$, for the 20 different values of z that lay in the interval between 0.0005 and 0.01.

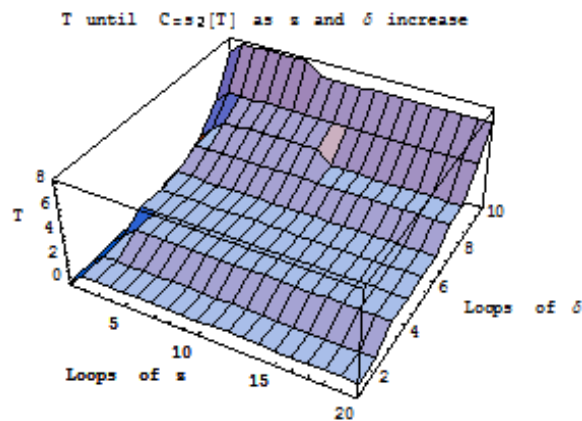


Figure 6: The simulated T that we derive by varying δ from .9 to .99 (in 10 steps of .01), and z from .0001 to .01 in 20 steps of .0005.

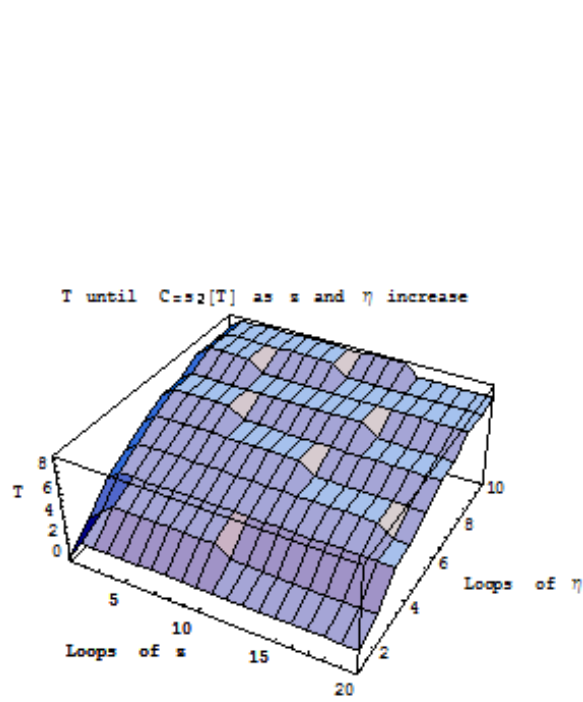


Figure 7: The simulated T that we derive by varying η from .1 to 1 (in 10 steps of .1), and z from .0001 to .01 in 20 steps of .0005.