INNOVATION AND IMITATION WITH AND WITHOUT INTELLECTUAL PROPERTY RIGHTS

RUFUS POLLOCK

UNIVERSITY OF CAMBRIDGE

JANUARY 2008

Abstract. An extensive empirical literature indicates that returns from innovation are appropriated primarily via mechanisms other than formal intellectual property rights – and that 'imitation' is itself a costly activity. However most theory assumes the pure nonrivalry of 'ideas' with its implication that, in the absence of intellectual property (for example under an 'open source' regime), innovation (and welfare) is zero. This paper introduces a formal model of innovation based on imperfect competition in which imitation is costly and an innovator has a first-mover advantage. Without intellectual property, a significant amount of innovation still occurs and welfare may actually be higher than with intellectual property.

Keywords: Innovation, Imperfect Competition, Intellectual Property, Imitation

JEL codes: K3, L5, O3

Emmanuel College, Cambridge University. Email: rp240@cam.ac.uk or rufus@rufuspollock.org. First Version: September 2006. My thanks for comments go to participants in seminars at Cambridge and Warwick as well as at the December 2006 DIME workshop in Manchester.

1. Introduction

Repeated surveys, such as Levin et al. (1987), Mansfield (1985), Cohen, Nelson, and Walsh (2000), and Arundel (2001), show that firms appropriate returns from innovation using a variety of methods including secrecy, lead time, marketing and sales, learning curve advantages and patents. Furthermore, they also suggest that for most industries (with a few notable exceptions) patent protection is of low importance. As Hall (2003) summarizes (p. 9): 'In both the United States and Europe, firms rate superior sales and service, lead time, and secrecy as far more important than patents in securing the returns to innovation. Patents are usually reported to be important primarily for blocking and defensive purposes.'

Of particular interest is the finding that imitation is a costly process both in terms of time and money, and one, furthermore, upon which the effect of a patent – if it has any effect at all – is to increase its cost not to halt it entirely. Perhaps most striking in this respect are Tables 8 (p. 810) and 9 (p. 811) of Levin et al. (1987) which summarize, respectively, reported cost of imitation (as a percentage of innovator's R&D expenditure) and time to imitate. For example, of the processes surveyed which were not protected by patents fully 88% had an imitation cost which was more than 50% of the innovator's initial outlay. For major products the analogous figure was 86%. Imitation also takes time: 84% of unpatented processes took 1 year of longer to imitate, while for products the analogous figure 82%.[1]

Such results indicate that for many innovations, even without patent protection, imitation involves substantial cost and delay.[2] Given this, as well as the strong impact the assumption of costless imitation has on our conclusions, it would seem important to investigate the consequences of weakening this presumption and, in particular, the possibilities of innovation without intellectual property rights.

However much of the existing theoretical literature has tended to assume 'perfect' non-rivalry, that is, that an innovation (or creative work) once made may be costlessly, and instantaneously, reproduced. The assumption is most often evident in the claim, which

[1] Of course, one must be cautious in interpreting such figures given the likely selection bias in deciding whether to patent or not – it is precisely those innovations which are hard to imitate without a patent which will not be patented.

[2] As emphasized by Dosi (1988) the distinction between innovation and imitation is often highly blurred and that imitation itself is a creative process.

follows directly from it, namely that without the provision of intellectual property rights such as patents and copyrights no innovation would be possible.

For example, Nordhaus (1969) (and following him Scherer (1972)), in what is considered to be one of the founding papers of the policy literature, implicitly assume that without a patent an innovator gains no remuneration. Similarly, Klemperer (1990) in his paper on patent breadth makes clear his assumption of costless imitation[3](p. 117): 'For simplicity, I assume free entry into the industry subject to the noninfringement of the patent and that knowledge of the innovation allows competitors' products to be produced *without* fixed costs and at the same constant marginal costs as the patentholder's product. Without further loss of generality, I assume marginal costs to be zero.' (Emphasis added). Many similar examples can easily be supplied in which imitation without intellectual property rights is implicitly, or explicitly, assumed to be 'trivial'.[4]

This paper, by contrast, provides a simple theoretical model in which costly imitation is central. Combined with first-mover advantage for the innovator we show that a significant amount of innovation takes place in the absence of intellectual property rights – even when imitation is cheaper than innovation. In addition we provide an easy and intuitive way to conceptualize, and model, the overall space of innovations which allows us to compare in a straightforward manner the relative performance of regimes with and without intellectual property rights, both in terms of innovation and welfare. This approach supplies several novel insights.

First, that as innovation costs fall 'allowable' imitation costs (that is imitation costs that still result in innovation being made) fall even faster. Thus, if the cost of innovation (relative to market size) differs between industries, then, even if relative imitation costs are the same, there will be very substantial difference in the impact of intellectual property rights. In particular, in the industry with lower innovation costs the gains for innovation and welfare with intellectual property rights will be much lower (and for welfare could

[3]Though it should be noted that it is possible to interpret the travel cost incurred by consumers in Klemperer's model as some form of 'design-around' or imitation cost that must be paid by competing firms. Nevertheless, in Klemperer's model, absent IP the innovator's gross profits (excluding the sunk cost of innovation) will be driven to zero by competition. As a result, anticipating a net loss, an innovator would not enter.

[4]See e.g. Scotchmer and Green (1990); Hopenhayn and Mitchell (2001); Menell and Scotchmer (2005).

even be negative).[5] As such, a main point of this paper is to show how the impact (and benefits/costs) of intellectual property rights may vary in a systematic way across industries. In particular there will be industries in which intellectual property rights are necessary – and industries where they are not – and this paper presents one basis for a taxonomy to sort out which is which.

Second, and relatedly, comparing regimes without and with intellectual property rights we show that the welfare ratio is systematically higher than the innovation ratio.[6] Moreover, this is not simply for the well-known reason that (conditional on the innovation occurring) without intellectual property rights greater competition results in increased output and lower deadweight losses. Rather, there is an additional factor, namely that the set of innovations occurring under an IP regime are, on average, less socially valuable because they have higher fixed costs of creation. Specifically, the model allows us to clearly distinguish three sources of welfare differences between the two regimes: first, less innovation occurs without intellectual property rights; second, the welfare of a given innovation is higher under competition that under monopoly; third, as just mentioned, innovations which occur only under an intellectual property regime are less valuable.

In addition to its 'stand-alone' uses, we also believe our model is valuable in its potential for integration into other innovation frameworks. In this paper, at least in relation to innovation, there is no downside to intellectual property rights and therefore, almost by assumption, an IP regime will outperform a no IP regime.[7] It would therefore be interesting to combine what we have here with more sophisticated models of the innovation process, for example one which incorporates cumulativeness. One of the main deficiencies of the cumulative innovation literature has been a lack of attention to the question of competition

[5]Consider, for example, pharmaceuticals compared to software. Starting a pharmaceutical (or biotech) company requires very substantial investment on the order of millions of euros while a software startup may need only a few tens of thousands of euros.

[6]The innovation ratio is the innovation level without intellectual property rights versus the level with intellectual property rights. Similarly the welfare ratios is the level of welfare without intellectual property rights versus the level with.

[7]Rather what we are trying to investigate here is how wide the gap is. With perfect nonrivalry without intellectual property rights innovation is zero. We show that allowing for non-zero imitation, even if quite small, can dramatically change this result.

in the end product market – and how such competition changes with the IP regime.[8] Combining this paper's explicit modelling of imitation and competition in the end product market with a more sophisticated model of innovation would deliver a 'best-of-both-worlds' model, with an improved ability to capture both the benefits, and costs, of intellectual property rights.

1.1. **Existing Literature.** There are, of course, some papers in the existing literature which do allow for non-trivial imitation. For example Gallini (1992), allows patented innovations to be imitated for some fixed cost K. With free entry of imitators, K is then the maximum income achieved by an innovator who patents. Thus, in this model, imitation costs must be higher than innovation costs for innovation to occur.[9] In our model, by contrast, imitation costs, both with and without intellectual property rights, may take any value (and without intellectual property rights are usually assumed to be less than innovation costs).

Other approaches include those based on locational models such as Waterson (1990) and Harter (1994) which both feature entry by a competing (imitative) firm within a horizontal product differentiation framework and focus on the impact of patent breadth on innovation and welfare.[10] This locational approach is obviously well-suited to considering imitation but is limited by the fact that it is extremely hard to endogenize entry. Both of the papers mentioned limit (imitative) entry to at most one firm. This makes it hard to analyze how changes in imitation cost impact on market structure and the innovator's rents. By contrast, we adopt a Stackelberg model of first-mover advantage. While this is obviously restrictive in other ways it allows us to tractably analyze equilibrium imitative entry.

Finally, Pepall and Richards (1994) also present a model which permits non-trivial imitation. Similar to our paper their model features Stackelberg competition with the

[8]For example, Bessen and Maskin (2006) assume in their model of cumulative innovation that, without intellectual property rights, each of the two firms receives some exogenously given share s of profits of that obtained with intellectual property rights. Meanwhile, Pollock (2006), following the approach of e.g. Denicolo (2000) and Bessen (2004) assumes that the IP regime only affects licensing and does not impact on the stand-alone value of the innovations.

[9]This is not precisely correct since Gallini allows for a firm not to patent – with non-patented inventions imitated at zero cost but only with some exogenous probability p_D. However, in this case (i.e. a firm does not patent) (a) there is no imitation cost – imitation either happens or it does not with some exogenous probability; (b) IP rights are irrelevant.

[10]The impacts of patents is rather different in the two models. In Waterson (1990) it is an exclusion zone enforced via imperfect litigation (with fixed imitation costs) while in Harter (1994) the effect of a patent has a rather different dual effect: it makes imitation cheaper but the imitator must locate her product outside of the exclusion zone set by the patent.

innovator taking the role of the leader. However, their focus is on quality choice by the innovator and how imitation may lead to welfare losses due to inefficiently low choice of product quality. We, on the other hand, are more interested in exploring how variations in relative imitation cost impact on innovation, and how, incorporated into a model of the distribution of innovations at the aggregate level, this in turn can be used to examine the relative welfare performance of different regimes.

2. The Model

As should be clear from the survey of the empirical data above, in modelling imitation there are two basic directions in which to advance: imitation may be costly in terms of money or in terms of time.[11]

Here we shall confine ourselves to the case of imitation which is costly in terms of money and shall retain the assumption that it is costless in terms of time, i.e. instantaneous. Specifically, we adopt a model based on the Stackelberg model of quantity competition with multiple followers.[12] In our case, the first mover role is naturally taken by the developer of the original innovation whom we term the 'innovator', and the role of followers by 'imitators'. In the Stackelberg game the first mover advantage derives from the ability to commit to a particular output level before other players. Here, however, it is better to see the Stackelberg framework simply as a convenient method for modelling an advantage that derives from far more general sources, for example lead time, learning curve effects and the ability to put in place a marketing and sales operation (to take some of the items frequently cited in the empirical literature referred to in the introduction).

In all other respects firms are the same except for the fact that the innovator has different fixed costs from those of imitators. These fixed costs, both of the innovator and the imitators, are assumed to be non-zero – this along with the first-mover advantage is the key aspect of the model and again this assumption is based on the empirical evidence that was discussed above. There is no formal delay in innovation but the Stackelberg framework implicitly assumes the first-mover has time enough to commit to supply as

[11]There are clearly other possibilities, for example imitation may be limited by the availability of skilled labour, or access to other necessary complementary assets (see e.g. Teece (1986)). However, these are both more complex to model and, we believe, of lesser importance than the main factors of time and money.

[12]It could therefore be argued there is some temporal aspect in that the innovator is able to 'move' before imitators. However, there is no real imitation lag in the sense of a period of time in which the original innovator enjoys a monopoly of the relevant market.

much of the market as she wishes. Demand is taken to be linear with an inverse demand curve $p(q) = a - bq$. To summarize:

(1) F_i the fixed cost of development for the innovator.
(2) F_m the fixed cost of imitation which is assumed to be common across all imitators. Also define ϕ to be imitation cost as a proportion of innovation cost, so $\phi = F_m/F_i$. We assume that imitation cost is always less than innovation cost and that in the presence of intellectual property rights imitation does not occur (which could be interpreted as having infinite imitation cost).[13]
(3) $c(q)$, marginal cost of production once the product is developed. It is assumed to be common between imitators and innovators (they both end up using the same technology), to be constant, and, without loss of generality, to be equal to zero.
(4) Linear demand given by $p(q) = a - bq$

We have a slight variation on the classic two-stage model in which the sequence of actions can be considered as falling into three periods as follows:

(1) An innovator decides whether to enter. If the innovator does enter then (s)he incurs a fixed cost, F_i, and develops a new product
(2) Imitators decide whether to enter. If an imitator does enter (s)he incurs a fixed cost of F_m, and then has capacity to produce the new product.
(3) Production occurs with price and quantities determined by Stackelberg competition in which the 'innovator' has the first-mover role and all imitators move simultaneously.

2.1. **A Normalization.** Define $k = \frac{a^2}{4b}$ so k is equal to half the area under the demand curve and therefore the level of monopoly profit. No agent's profits (innovator or imitator) can be greater than monopoly profits k. Hence let us simplify by normalizing all profits and fixed costs by dividing them by k – equivalent to setting k equal to 1 in the analysis below. Thus from now on when profits or fixed costs are discussed they should be taken not as absolute levels but as proportions of monopoly profits (itself equal to half of total potential welfare offered by the innovation). Formally:

[13]Note that this does not fit with the empirical data from Levin et al. (1987) where in several cases the costs of imitation exceeded those incurred by the innovator. Nevertheless, as the assumption greatly simplifies the analysis and incorporating the more complex reality would only strengthen our results, we feel warranted in proceeding as indicated.

$$f_i = \frac{F_i}{k}$$

$$f_m = \frac{F_m}{k}$$

Note that, ϕ, the ratio imitation cost is also equal to the ratio of the normalized costs: $\phi = F_m/F_i = f_m/f_i$.

2.2. **The Space of Innovations.** In this model an innovation is specified by the tuple consisting of its 'innovation' cost and its 'imitation' cost: (f_i, f_m) (or equivalently (f_i, ϕ).[14] Innovation and imitation costs are non-negative, $f_i, f_m > 0$, and we have assumed that imitation costs are never more than innovation costs: $f_m \leq f_i$. Furthermore, it will never be optimal for an innovator to enter if $f_i > 1$, since the maximum possible profits from entering the market (k) are less than the cost of the innovation.

Thus, under the assumptions given and using normalized variables the space of innovations is $IS = \{(f_i, f_m) \in [0,1] \times [0,1] : f_m \leq f_i\} = \{(f_i, \phi) \in [0,1] \times [0,1]\}$.

2.3. **Policy Regimes and the Effect of Intellectual Property Rights.** We will wish to consider different policy regimes. A given policy regime (R) has an associated model which will determines the costs and rents for the different agents and thereby defines some region in innovation space, IS, in which innovation occurs. It will also determine the welfare which each of those innovations generates.

In addition, a policy regime (R) will be taken to define a distribution of innovations over innovations over the innovation space IS which can be represented by some density function, say g^R. This function is primarily intended to capture information about the distribution of innovations at the aggregate level, for example industry or economy wide. This will be important because one cannot make decisions about the strength or presence of intellectual property rights on a firm-by-firm or technology-by-technology basis. Instead a policy-maker must set them at a very macro level – for example the length of patent protection is set by international treaty and must be the same across all patentable technologies. Even where there is choice, as in recent debates as to whether to extend

[14]This conveniently allows us to visualize innovation space in a two dimensional graph (see the figures below for examples).

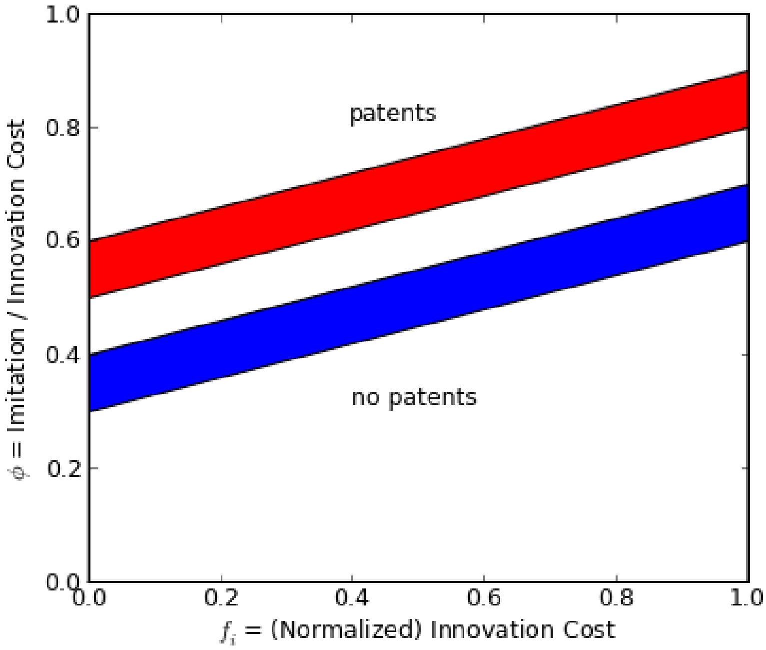

FIGURE 1. The Effect of Different Policy Regimes under the 'Breadth' approach. The light-shaded region above the diagonal indicates the area in which innovation occurs (the IP and no IP regime share a common model so this region is the same for both regimes). The lower band labelled 'no patents' indicated the distribution of innovations under the no IP regime while the upper band (labelled 'patents') indicates the new distribution of innovation with intellectual property rights (the implicit assumption here is that the introduction of intellectual property rights raises imitation costs by some fixed proportion leaving innovation costs unchanged).

patentability to software or copyright to perfumes, the decision must be made for an entire class of products displaying very substantial heterogeneity.[15]

In this paper we shall be interested in comparing and contrasting two particular regimes: that with intellectual property rights (e.g. patent or copyright) and that without. As just discussed, these regimes can differ both in their model (which determines whether a given innovation (f_i, f_m) occurs and the welfare it generates) and in the distribution of innovations over innovation space.[16] We focus on two distinct possibilities, with the first approach being the one we shall use by default:

[15] A secondary purpose for the distribution function is to capture uncertainty by interpreting this function as representing the 'beliefs' of a policy-maker.

[16] In some ways allowing variation in the distribution of innovations is redundant since any variation in distribution could be incorporated as a difference in models. However, changes in distributions provide a simpler approach, that is less cumbersome in notation and more intuitive for understanding.

(1) Models differ, distributions the same. Specifically, under the no IP regime we use the Stackelberg model presented above. With IP we assume that all imitation is prohibited and that, as a result, the innovator makes monopoly profits.[17]

(2) Models the same, distributions differ. Specifically, both regimes use the 'Stackelberg' model presented above but the distribution of innovations under no IP, g, is transformed to a new distribution g' under the IP regime. A graphical illustration of what this means is presented in Figure 1.

For future reference we shall label the first case the 'Zero Imitation' (ZI) approach to modelling intellectual property rights (and label the associated regime the 'Zero Imitation' regime) and the second the 'Breadth' (BR) approach to modelling intellectual property rights.

3. Solving the Model

We solve by recursing backwards through the game. First, in Proposition 1, we determine the solution to the Stackelberg model of price competition in the final product market assuming a fixed, given number of imitators. Next we solve for the equilibrium number of imitators using the zero-profit condition generated by the assumption of free entry. This gives the number of imitators as a function of the imitation cost f_m. Using this, we can determine the innovator's expected gross profits as a function of the number of imitators (and hence imitation cost f_m). If these profits exceed the innovation cost, f_i then the innovator would enter and the innovation occurs – otherwise it does not. We summarize the results in Propositions 3 and 4, which details the set of innovations occurring in equilibrium.

Proposition 1. *Let n be the exogenously given number of imitators. The solution to the Stackelberg model of competition by quantify is as follows where k is defined as above to equal $a^2/4b$ ('i' subscripts are on variables related to the innovator and 'm' subscripts are on variables related to an imitator):*

[17]This can be nested within our 'Stackelberg' model by restricting the number of imitators to be 0.

$$\begin{aligned}
q_i &= \frac{a}{2b} \\
q_m &= \frac{a}{2b(n+1)} \\
\textit{Total output} &= Q = \frac{a(2n+1)}{2b(n+1)} \\
p &= a - bQ = \frac{a}{2(n+1)} \\
\textit{Gross profits of an innovator} &= \Pi_i = \frac{k}{n+1} \\
\textit{Gross profits for an imitator} &= \Pi_m = \frac{k}{(n+1)^2} = \frac{\Pi_i}{n+1}
\end{aligned}$$

Proof. Omitted (the solution to the Stackelberg model is well-known). □

Proposition 2. *Imposing a zero net profit condition on the basis of free entry as an imitator, the number of imitators, n^e is as follows:*

- *Non-integer n:* $n^e = \sqrt{\frac{1}{f_m}} - 1$
- *Integer n:* $n^e = \max\{n \in \mathbb{Z} : f_m \leq \frac{1}{(n+1)^2}\}$

Proof. Allowing non-integer n we solve $\Pi_m = F_m$. This gives:

$$n^e = \sqrt{\frac{k}{F_m}} - 1 = \sqrt{\frac{1}{f_m}} - 1$$

Restricting to integer n we require the n such that $\Pi_m \geq F_m$ but with $n+1$ imitators $\Pi_m < F_m$. Substituting for Π_m gives the condition. □

Proposition 3. *Allowing the number of imitators to take non-integer values then an innovation (f_i, f_m) occurs if $f_m \geq f_i^2$ $(\phi \geq f_i)$. Thus, the set of innovations which occur is given by:*

$$A^c = \{(f_i, f_m) \in IS : f_m \geq f_i^2\} = \{(f_i, \phi) \in IS : \phi \geq f_i\}$$

Proof. Innovation only occurs if expected (net) profits are positive, that is $\Pi_i \geq f_i$. Substituting for the LHS using our value for the number of imitators from Proposition 2 gives the condition:

$$f_m \geq f_i^2$$

□

Proposition 4. *Restricting the number of imitators to integer values the set of innovations that occur is:*

$$A^{int} = \cup_{n=0}^{\infty}\{(f_i, f_m) \in IS : \frac{1}{n^2} \geq f_m > \frac{1}{(n+1)^2}, f_i \leq \frac{1}{n+1}\}$$

Proof. Direct from Proposition 2 □

Remark 1. Note the substantial difference between the two situations (non-integer and integer numbers of imitators). For example, with integer-only number of imitators, $f_m \geq \frac{1}{4} \Rightarrow n = 0$ and all innovations with $f_i \leq 1$ are realized, a very different outcome to that with continuous number of imitators. We return to this theme below, in Proposition 6.

In this model an innovation is defined by a pair (f_i, f_m) giving its innovation and imitation cost. We can therefore visualise potential innovations in a two dimensional graph of innovation/imitation cost space. In particular, we can summarize the results of the previous propositions in Figure 2. In this diagram the light-shaded (yellow) region is that in which innovations occur with non-integer numbers of imitators permitted, while the innovations in the dark-shaded (red) and light-shaded region occur when restricting to integer numbers of imitators. (The region above the diagonal should be ignored since we are assuming that imitation cost is always less than innovation cost).

While the preceding diagram is entirely correct as it stands, it will be useful to visualize the same data in a slightly different manner. We do this by replacing imitation cost by 'proportional' imitation cost (ϕ) – i.e. imitation cost as a proportion of innovation cost. Under our assumption that imitation cost is always less that innovation cost this means that we now have a constant range, [0,1], for 'proportional' imitation cost at all levels of innovation cost and, in visual terms, we have a uniform level of innovation per unit of innovation cost. This is shown in Figure 3 which is simply a re-rendering of Figure 2 using proportional innovation cost.

Proposition 5. *With intellectual property rights (zero imitation) all innovations in IS occur and $A^{IP} = IS$*

Proof. We have assumed that with intellectual property rights no imitation is possible hence an innovation occurs if and only if innovation costs are less than 1. □

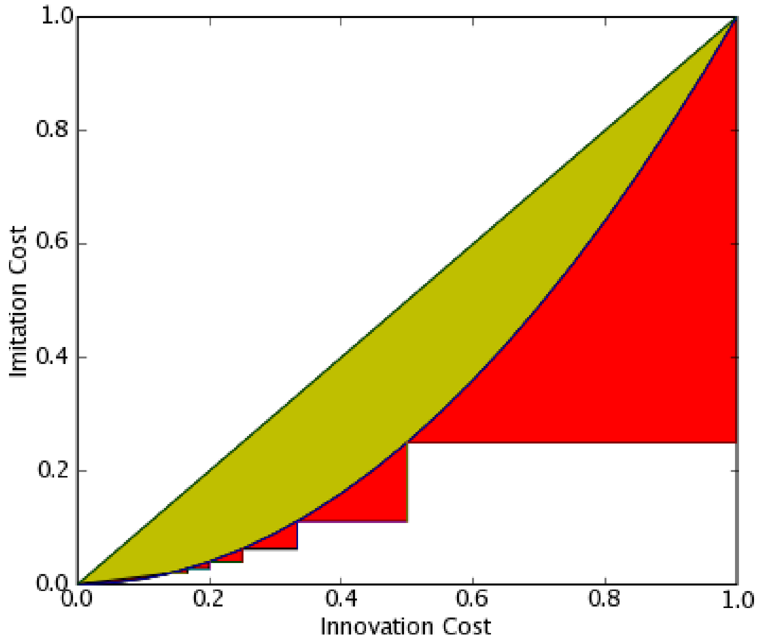

Figure 2. Innovations which occur without intellectual property rights (light shaded region: non-integer numbers of imitators allowed; dark-shaded extra innovations occurring when restricted to integer numbers of imitators).

Thus, with IP, all of the area under the 45 degree line in Figure 2 and all of the area in Figure 3 would be shaded.

Returning to our theme of the difference between allowing continuous and integer numbers of imitators, we have:

Proposition 6. *Assuming a uniform distribution over the space of innovations shown in Figure 2 (this corresponds to calculating area), that is with density function $g(f_i, \phi) = 1$, the ratio of innovation without intellectual property rights to that with intellectual property rights is: 50% (non-integer n), 72% (integer n).*

Proof. See appendix. □

Thus restricting to integer n increases the amount of innovation by nearly 50% and much of this extra innovation occurs at the higher levels of innovation and imitation cost when the number of imitators in the integer case will be low (zero, one or two). Despite, this difference in the remainder of the paper we shall, by default, focus on the case of continuous n. This is for two reasons. First, especially when performing integrations to

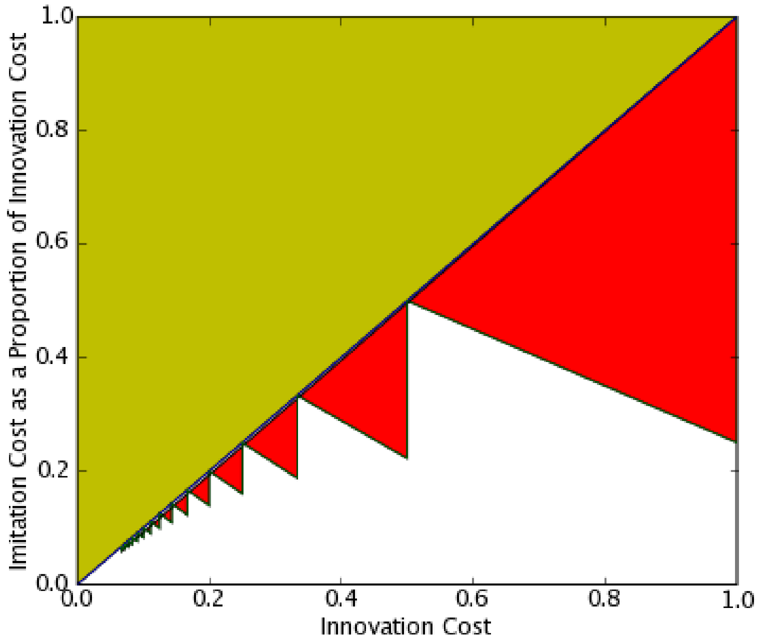

Figure 3. Innovations which occur without intellectual property rights (light shaded region: non-integer numbers of imitators allowed; dark-shaded extra innovations occurring when restricted to integer numbers of imitators).

obtain welfare totals, the continuous case is much easier to use. Second, as just shown, restricting to integer n will only strengthen our results regarding the relative performance of a no IP regime. Thus, any result we obtain for continuous numbers of imitators, will hold a fortiori for discrete number of imitators.

4. Welfare and Policy

From a policy perspective what really matters is the utility generated by innovation not how much innovation occurs. If the welfare from innovations realized without intellectual property rights differ systematically from those that are not or the welfare generated by a given innovation differs under the two regimes then welfare outcomes will differ from innovation levels.

Let R and S denote two distinct policy regimes. Define:[18]

$$\begin{aligned} W^R(f_i,\phi) &= \text{Welfare under regime R from innovation}(f_i,\phi) \\ \Delta W_S^R(f_i,\phi) &= W^R(f_i,\phi) - W^S(f_i,\phi) \end{aligned}$$

4.1. **Welfare Per Innovation.** Take R to be the no IP regime (NIP) and S to be the IP/zero imitation regime (ZI). Recall that under ZI all innovation in the innovation space, IS, occur. Let A denote the region in which innovation occurs under NIP, then we have:

Proposition 7. *The difference in welfare generated by an innovation (f_i,ϕ) under the no IP regime (NIP) compared to the zero imitation regime (ZI) is:*

$$\Delta W^{NIP}(f_i,f_m) = \begin{cases} \frac{n^{e2}}{2(n^e+1)^2}, & (f_i,\phi) \in A \\ -W^{ZI}(f_i,\phi), & (f_i,\phi) \in IS - A \end{cases}$$

In particular, when the innovation is in A – and therefore occurs under both regimes – this difference is always non-negative and the no IP regime generates more welfare than the zero imitation regime.

Proof. See appendix. □

The ΔW term captures the fact that, for a *given innovation*, the welfare generated by it differs between the two regimes. This difference is driven by two distinct, and contrary, effects. First, no intellectual property rights leads to greater competition. This transfers rents from producers to consumers and reduces the deadweight loss because total output expands. Second, with imitation there is greater entry which means total fixed costs expended for a given innovation are higher due to the greater number of producers. In this model, the first effect outweighs the second (conditional, of course, on the innovation still being produced without intellectual property rights).[19]

4.2. **A Single Technology With Observable Costs.**

Corollary 8. *Assume costs are precisely observable by a regulator. If IP is represented by the 'Zero Imitation' regime, the optimal policy rule is to grant intellectual property rights*

[18]Note that if the innovation (f_i,ϕ) does not occur under regime R then $W^R(f_i,\phi) = 0$.

[19]This result has a simple, intuitive, basis. Under a Stackelberg model of quantity competition the output of the leader (the innovator) stays fixed at the monopoly level. Thus, the income used to cover imitators' fixed costs must always come from output expansion. Hence, though imitative entry does increase fixed costs those fixed costs are always less than the increase in surplus arising from the output expansion.

if and only if the square of innovation costs (as a proportion of monopoly profits) is larger than imitation costs (also as a proportion of monopoly profits), that is: $f_i^2 > f_m$.

Proof. Our previous result shows that welfare without intellectual property rights is greater than than with intellectual property rights (Zero Imitation) if and only the innovation occurs without intellectual property rights. Thus the 'square' rule follows directly from from our result on innovation as described in Propositions 2 and 3. □

This 'square' rule is we believe a novel result in the literature. While its convenient form is clearly specific to the Stackelberg-type model we have adopted, as we show below, the point that the 'allowable' imitation cost falls (that is the minimal imitation cost such that innovation still occurs) as innovation cost falls is a general one.

We also note that if IP is represented by the 'Breadth' regime rather than a 'Zero Imitation' regime a very similar result still obtains. To be precise, assuming that an increase in breadth acts to increase imitation costs leaving innovation costs unchanged, then, given an innovation with costs (f_i, f_m) (under no IP), the optimal policy rule consists in setting the breadth of the IP right such that if f'_m is the new imitation cost (under IP) then $f'_m = f_i^2$.

4.3. **A Distribution of Innovations.** The results of the previous section are certainly valuable, however, they suffer from two significant drawbacks if intended for use by regulators in the real-world. First a regulator usually lacks precise information about innovation and imitation costs (at least ex-ante). Second, and more importantly, as discussed above in Section 2.3, a policy-maker cannot make decisions about the strength or presence of intellectual property rights on a technology-by-technology basis. Instead decisions about the existence, and strength, of such rights must be taken at a much more aggregate level.[20]

Thus, in this section we extend our welfare analysis to the aggregate, industry or economy-wide, level by incorporating the distribution of innovations. Using the notation set out in the Section 2.3 above we encapsulate the distribution of innovations under a given regime, R, in a probability distribution function g^R defined over the space of innovations IS. Extending our existing notation we have:

[20]And this is not simply for informational reasons but because of the need to be compatible with existing norms and agreements. For example, an international treaty (TRIPS) sets down a minimum length for patent protection and mandates that it must be the same across all patentable technologies.

$$\begin{aligned} W^R(X) &= \text{Welfare from region X under regime R} = \int_X W^R(f_i, f_m) g^R \\ \Delta W_S^R(X) &= W^R(X) - W^S(X) \end{aligned}$$

We shall focus again on the no IP (NIP) and zero imitation (ZI) IP regime. As stated in Section 2.3, we assume these share the same distribution of innovations. We shall therefore drop the superscript and simply use g for this distribution. Recall also that, under the zero imitation regime, all innovation in IS takes place. Let A be the region in which innovation takes place under no IP and define $B = IS - A$, that is, the set of innovations not in A. Then:

$$W^{ZI} = W^A(ZI) + W^{ZI}(B) \tag{4.1}$$

$$W^R = W^R(A) + W^R(B) = W^{ZI}(A) + \Delta W_{ZI}^R(A) \tag{4.2}$$

The second equation illustrates how we may break up the welfare under regime R. First, note that the welfare from region B, $W^R(B)$ is zero since, by definition, no innovation occurs in that region. Turning to region A, we may divide welfare that we would get in the case of zero imitation (the first term) plus the difference between that level and the level of welfare in regime R: ΔW.

This allows us to distinguish between three effects that operate with respect to differences in welfare. First, less innovation occurs under no IP compared to Zero Imitation. Second, is the fact, already mentioned, that, assuming an innovation occurs under both regimes, it will generate more welfare under no IP than under Zero Imitation. This is captured in the ΔW term. Third, is the fact that innovation fixed costs may differ systematically between regions A and B (A is the region in which innovation occurs under both regimes while B is everything else). This will materialize in the relative sizes of $W(A)$ and $W(B)$. We illustrate these effects with a simple example where innovations are uniformly distributed:

Proposition 9. *Assuming a uniform distribution over the space of innovations as shown in Figure 2, that is with density function $g(f_i, \phi) = 1$, welfare levels are as follows (where*

NIP indicates a regime without intellectual property rights and the number of imitators may take non-integer values):

$$W^{ZI}(A) = \tfrac{7}{12}, \textit{average welfare density} = \tfrac{7}{6} \tag{4.3}$$

$$W^{ZI}(B) = \tfrac{5}{12}, \textit{average welfare density} = \tfrac{5}{6} \tag{4.4}$$

$$\Delta W(A)^{NIP} \approx \tfrac{2}{12}, \textit{average welfare density} = \tfrac{2}{6} \tag{4.5}$$

Proof. See appendix. □

Thus, the ratio of welfare without intellectual property rights to a situation in which they are present is **75%**. Comparing this with the results of Proposition 6 we see that a regime without intellectual property rights while only having half the level of innovation delivers three quarters of the welfare achieved with intellectual property rights. Furthermore, we see that the third effect mentioned above, that is the systematic difference in the fixed cost of innovation, is a significant driver of these results. For example, if we were to assume that ΔW were zero, that is the welfare generated by innovations under no IP and IP were the same, we would still have a welfare ratio of 58% – the same gain if there under the converse assumption of no difference in fixed costs but only differences in per innovation welfare yields.

To give another illustration, consider now the question of uncertainty. Suppose a policy-maker knows precisely the proportional imitation costs but has complete uncertainty regarding innovation costs (so the policy-makers belief are represented by a uniform distribution over the possible values).[21]

Proposition 10. *Assuming a uniform distribution of innovation costs if imitation costs are more than 70% of innovation costs then welfare is higher without intellectual property rights.*

Proof. See appendix. □

Turning to the case where innovation costs are known with certainty but there is complete uncertainty regarding imitation costs one has a similar result:

[21]For example, the data provided in Levin et al. (1987) provide information on proportional imitation costs but nothing on the cost of innovation itself.

Proposition 11. *Assuming a uniform distribution of proportional imitation costs, if innovation costs are less than 20% of total potential monopoly profits then welfare is higher without intellectual property rights.*

Proof. See appendix. □

5. The General Case

The quantitative results obtained above must clearly be specific to assumptions regarding the underlying model and distribution of innovations. However, the basic point that welfare proportions will always be systematically higher than innovation proportions (even if we ignore deadweight loss) holds in general.

Recall that an innovation is specified by the tuple (f_i, f_m) (or equivalently (f_i, ϕ) and that (using normalized variables) the space of innovations is then $IS = \{(f_i, f_m) \in [0,1] \times [0,1] : f_m \leq f_i\} = \{(f_i, \phi) \in [0,1] \times [0,1]\}$.

Now, any given regime R (with associated model of innovation and imitation M^R) will define some region in IS in which innovation occurs. Following previous convention we will denote this region by A. We make the mild assumptions that:

Assumption 12. Suppose the innovation $I^1 = (f_i^1, f_m^1) \in A$ then:

(1) Any other innovation with the same imitation cost but lower innovation cost occurs under R. Formally: $\forall f_i \leq f_i^1, (f_i, f_m^1) \in A$.
(2) Any other innovation with the same innovation cost but higher imitation cost occurs under R. Formally: $\forall f_m \geq f_m^1, (f_i^1, f_m) \in A$.

How can we characterise this region, A, in which innovation occurs under regime R? Define $h(f_i)$ as the infinum of all innovations with innovation cost f_i that are in A:

$$h(f_i) = inf\{f_m : (f_i, f_m) \in A\}$$

Let us assume (without loss of generality) that $h(f_i) \in A$.

Proposition 13. *The area in which innovation occurs A is given as follows:*

$$A = \{(f_i, f_m) \in IS : f_m \geq h(f_i)\}$$

Furthermore, h is a non-decreasing function.

Proof. The first part follows directly from Assumption 12.2 combined with the definition of the supremum h. To show that h is non-decreasing suppose not, that is that there exists $f_i^1 < f_i^2$ such that $f_m^1 = h(f_i^1) > h(f_i^2) = f_m^2$. By Assumption 12.1 $(f_i, f_m^2) \in A, \forall f_i < f_i^2$ which implies, in particular, $(f_i^1, f_m^2) \in A$, but $f_m^2 < f_m^1$ which implies $h(f_i^1) \leq f_m^2 < f_m^1 = h(f_i^1)$ which is a contradiction. □

Definition 14. Given a regime R recall that I^R is the amount of innovation occurring under R and W^R the total amount of welfare. Then given two different regimes, R, S, define:

(1) IR(R,S) = Innovation Ratio of R to S = the ratio of innovation under R to innovation under S
(2) WR(R,S) = Welfare Ratio of R to S = the ratio of welfare under R to welfare under S

Proposition 15 (Welfare Ratio is higher than Innovation Ratio)**.** *Take a general regime R and a corresponding zero imitation (ZI) regime (so the ZI regime shares the same distribution of innovations as R). Assume that welfare from a given innovation (if it occurs under both regimes) generates at least as much welfare under R as under ZI:*

$$W^R(f_i, f_m) \geq W^{ZI}(f_i, f_m)$$

Then the welfare ratio *of R compared to zero imitation ZI will be greater than or equal to the* innovation ratio *of R compared to zero imitation (ZI). Furthermore, the inequality is strict if there is any innovation which occurs under R and there are some innovations which occur under ZI but not under R. That is:*

$$WR(R, ZI) \geq IR(R, ZI)$$

Proof. See appendix. □

Remark 2. Note that this result holds even if there are *no* deadweight losses, that is the welfare generated under R per innovation is the same as under ZI. Hence, this proposition establishes in great generality the point made earlier that the narrowing of the differential between the no IP and IP regime when moving from innovation to welfare was driven not simply by the well-known welfare-benefits of greater competition but also by systematic differences in the average of costs of innovations occurring with and without IP.

6. Conclusion

In this paper we have presented a simple model of innovation with imitation. We have shown that when imitation is costly and there is some form of first mover advantage the initial innovator may still be able to garner sufficient rents to cover the fixed cost of development even though not enjoying a pure monopoly. As discussed in the introduction, there is a great deal of empirical support for believing imitation costs and first mover advantage are important. This paper demonstrates that these concerns can be analyzed simply and tractably, and, that doing so, generates important new insights – most significantly that ignoring them may overstate the importance of intellectual property rights.

Here innovations are specified by a tuple consisting of the 'innovation' cost and the 'imitation' cost (the innovation cost being the cost to the first developer of the product). Using our Stackelberg-based model of first-mover advantage we obtained a precise description of which innovations would occur with imitation (that is, without IP rights). The formula took a particularly simple form which we dubbed the 'square' rule because it stated that innovations occurred if and only if (normalized) imitation cost was greater than the square of (normalized) innovation cost (we normalized by dividing by the potential monopoly profit so that all costs were in the range $[0, 1]$). While this particular formula must necessarily be dependent on the precise structure of the underlying model, the basic point that 'allowable' imitation costs fall with innovation cost is, we believe, a very general one – one, furthermore, which has received scant notice in previous literature.

Next we turned to a consideration of welfare and its implication for policy. We first showed that the 'square' rule carried over from innovation to welfare. This has important policy consequences. For example, if the ratio of imitation costs to innovation costs are the same in two industries but the (normalized) cost of innovation differs, then the impact of intellectual property rights in the two industries will be very different. Specifically, in the industry with lower innovation costs, the benefits of IP will be much lower (and could even be negative). This result illustrates how the impact of IP may vary in a systematic way across industries. In particular there will be industries in which intellectual property rights are necessary – and industries where they are not, and this paper presents one basis for a taxonomy to determine which is which.

However, it is rare that a policy-maker knows precisely the innovation and imitation costs for a given technology. Furthermore, it is, in practice, impossible for a policy-maker to set the level of IP on a technology, or even industry-by-industry basis. Hence, the next step was to extend our analysis to consider the case where there is a *distribution* of innovations – this distribution can be taken to represent either beliefs, or a collection of potential innovations at the industry or economy-wide level.

Comparing regimes without and with intellectual property rights we showed that the welfare ratio is systematically higher than the innovation ratio. Moreover, it was demonstrated that this is not simply for the familiar reason that, conditional on the innovation being made, greater competition without intellectual property rights leads to in increased output and lower deadweight losses. Rather, there was the additional factor, namely that the set of innovations occurring under an IP regime are, on average, less socially valuable because they have higher fixed costs of creation.

Finally, we note that there are a variety of way in which the present work could be extended. One could, for example, introduce a 'race' for the innovation in standard manner. This would allow for multiple firms at the innovation stage competing to produce the original innovation. This could be extended so that failed innovators can be imitators at the second stage.

On a separate point, one distinctive feature of this model is that intellectual property rights always lead to maximal innovation. In a more complex model, for example one involving cumulative innovation, this might no longer be the case. There are a variety of approaches that could be taken to integrate such dynamics and investigating these options would be one of most important improvements to the model that could be made.

Another option, which has already been mentioned briefly, is to have a richer model of imitation delay. Similarly, allowing for types of imperfect competition other than Stackelberg would also be a valuable extension. For example, the models of Waterson (1990) and Klemperer (1990) both provide for product differentiation and these models could be adapted to provide a richer and more realistic model of imitation in the presence – and absence – of intellectual property rights.

A. Proofs of Propositions

Proof of Proposition 6. A uniform distribution of innovation corresponds to the standard euclidean measure over IS, which in turns corresponds to calculating areas in Figure 2. With intellectual property rights no imitation is permitted so all the innovations in the figure occur (total area of the figure is 1). Thus to calculate the proportions of innovation occurring without intellectual property rights we need to calculate the size of the dark-shaded and light-shaded areas as proportion of the entire figure.

For continuous n we consider the light-shaded region. This, clearly, has area equal to 1/2.

Restricting to integer n we need to add to this the area of the dark-shaded (red) region. The area of the dark-shaded (red) region is made up of a series of similar triangles. The nth triangle (working down from the largest) has area:

$$0.5 \cdot b \cdot h = 0.5 \cdot (\frac{1}{n} - \frac{1}{n+1}) \cdot (\frac{1}{n} - \frac{n}{(n+1)^2})$$

Thus total area of dark-shaded (red) region is:

$$0.5 \sum_{1}^{\infty} (\frac{1}{n} - \frac{1}{n+1}) \cdot (\frac{1}{n} - \frac{n}{(n+1)^2}) = 0.5 \cdot (\sum \frac{1}{n^2} - \sum \frac{1}{n(n+1)} - \sum \frac{1}{(n+1)^2} + \sum \frac{n}{(n+1)^3})$$

All of these sums are simple except for the third. For this one approximate as follows:

$$\sum \frac{n}{(n+1)^3} \approx \sum_{1}^{99} \frac{n}{n+1}^3 + \int_{99}^{infty} \frac{1}{(x+1)^2} = 0.432976 + 0.01 = 0.4430$$

Substituting this gives the dark-shaded (red) region's total area as:

$$0.5 \cdot (\sum \frac{1}{n^2} - \sum \frac{1}{n(n+1)} - \sum \frac{1}{(n+1)^2} + \sum \frac{n}{(n+1)^3}) = 0.5 \cdot ((1+X) - 1 - X + 0.4430) = 0.2215$$

Thus total area of light-shaded and dark-shaded region is $0.5 + 0.2215 \approx 0.72$.

□

Proof of Proposition 7. First let us determine the welfare arising from a given innovation. If there are n imitators we have that consumer surplus (CS) and producer surplus (PS) are as follows:

$$CS(f_i, f_m) \quad = 0.5 \cdot (a-p) \cdot q = \tfrac{(2n+1)^2}{2(n+1)^2} \tag{A.1}$$

$$PS(f_i, f_m) \quad = \Pi_i - f_i + n \cdot (\Pi_m - f_m) = \tfrac{1}{n+1} - f_i \tag{A.2}$$

Note that we have used the fact that, with continuous n, the zero profit condition implies $\Pi_m = f_m$. Summing to get total welfare we have:

$$W(f_i, f_m) = CS + PS = \frac{(2n+1)^2}{2(n+1)^2} + \frac{1}{n+1} - f_i$$

Now in a ZI regime $n = 0$ so:

$$W^{ZI} = \frac{3}{2} - f_i$$

Thus,

$$\Delta W(f_i, f_m) \quad = W^R(f_i, f_m) - W^{ZI}(f_i, f_m) \tag{A.3}$$

$$= (\tfrac{(2n+1)^2}{(n+1)^2} + \tfrac{1}{n+1} - f_i) - (\tfrac{3}{2} - f_i) \tag{A.4}$$

$$= \tfrac{n^2}{2(n+1)^2} \tag{A.5}$$

$\square$

Proof of Proposition 9. To calculate total welfare for region X we integrate welfare per innovation, $W(f_i, f_m)$, over X.

$$W^{ZI}(A) = \frac{1}{2}(\frac{3}{2} - \text{avg over A}(f_i)) = \frac{3}{4} - \frac{1}{2}\frac{1}{3} = \frac{7}{12}$$

$$W^{ZI}(B) = \frac{1}{2}(\frac{3}{2} - \text{avg over B}(f_i)) = \frac{3}{4} - \frac{1}{2}\frac{2}{3} = \frac{5}{12}$$

Calculating ΔW is slightly more complicated:

$$\Delta W(A) = \int_A \frac{n^2}{2(n+1)^2} = \int_0^1 \int_{f_i}^1 d\phi df_i$$

Recall that:

$$\phi = \frac{f_m}{f_i} \tag{A.6}$$

$$n+1 = \frac{1}{\sqrt{f_m}} \Rightarrow \frac{n^2}{(n+1)^2} = 1 - 2\sqrt{f_m} + f_m \tag{A.7}$$

Thus, substituting f_m for ϕ as well as for n we have:

$$\Delta W(A) = 0.5 \int_0^1 \frac{1}{f_i} \int_{f_i^2}^{f_i} 1 - 2\sqrt{f_m} + f_m df_m df_i$$

Working through the first integration gives:

$$\Delta W(A) = 0.5 \int_0^1 1 - \frac{4\sqrt{f_i}}{3} - \frac{f_i}{2} + \frac{4f_i^2}{3} - \frac{f_i^3}{2} df_i = \frac{13}{72} \approx \frac{1}{6}$$

□

Proof of Proposition 10. We need to determine welfare at a particular level of ϕ (imitation cost as a proportion of innovation costs) assuming a uniform distribution of innovation costs under an IP (zero imitation) and no IP regime. Proceeding as above but making all welfare calculations a function of ϕ we have:

$$W^{ZI}(A)(\phi) = \tfrac{1}{2}(3\phi - \phi^2) \tag{A.8}$$

$$W^{ZI}(B)(\phi) = 1 - \tfrac{1}{2}(3\phi - \phi^2) \tag{A.9}$$

$$\Delta W^{NIP}_{ZI}(A)(\phi) = \tfrac{1}{2}(\phi - \tfrac{4}{3}\phi^2 + \tfrac{\phi^3}{2}) \tag{A.10}$$

The difference in welfare between a regime without IP compared to one with is $\Delta W(\phi) = W^{NIP}(\phi) - W^{ZI}(\phi)$. Thus to determine the cut-off point, α say, such that for all $\phi \leq \alpha$ the no IP regime is preferable we simply need to solve:

$$\Delta W(\phi) = 0$$

(Note that ΔW is an increasing function of ϕ so the solution will be unique and that $\Delta W(0) < 0$ and $\Delta W(1) > 0$ so a solution will exist).

Proceeding numerically we obtain a figure of $\alpha = 0.704 \approx 0.7$. □

Proof of Proposition 11. We proceed as in the previous proof though this time focusing on welfare at a particular level of f_i (innovation cost as a proportion of potential monopoly profit) assuming a uniform distribution of proportional imitation cost under an IP (zero imitation) and no IP regime. Making all welfare calculations a function of f_i we have:

$$W^{ZI}(A)(f_i) = (\tfrac{3}{2} - f_i)(1 - f_i) \tag{A.11}$$

$$W^{ZI}(B)(f_i) = (\tfrac{3}{2} - f_i)f_i \tag{A.12}$$

$$\Delta W^{NIP}_{ZI}(A)(f_i) = \tfrac{1}{2}(1 - \tfrac{4\sqrt{f_i}}{3} - \tfrac{f_i}{2} + \tfrac{4f_i^2}{3} - \tfrac{f_i^3}{2} \tag{A.13}$$

The difference in welfare between a regime without IP compared to one with is $\Delta W(f_i) = W^{NIP}(f_i) - W^{ZI}(f_i)$. Thus to determine the cut-off point, α say, such that for all $f_i \leq \alpha$ the no IP regime is preferable we simply need to solve:

$$\Delta W(f_i) = 0$$

(Note that ΔW is a decreasing function of f_i so the solution will be unique and that $\Delta W(0) > 0$ and $\Delta W(1) < 0$ so a solution will exist).

Proceeding numerically we obtain a figure of $\alpha = 0.191 \approx 0.2$. □

Proof of Proposition 15. Claim: Assume the innovation $(f_i^1, f_m^1) \in A$. Then for any regime X if $f_i < f_i^1, W^X(f_i^1, f_m^1) > W^X(f_i, f_m^1)$.

Proof of Claim. Innovation cost is a sunk cost and the original innovation (f_i^1, f_m^1) is in A (and so occurs under either regime). Then reducing the cost of innovation has no effect on the behaviour of the innovator and as imitation cost are unchanged the solution of the model in terms of price, output etc must be the same. As a result Consumer Surplus must be unchanged and the only change to producer surplus comes from a reduction in the innovator's cost (which increases producer surplus). The claim follows. □

Under ZI all innovations in IS occur. Let A be the region of IS in which innovations occur under R. Let g be the probability distribution function on IS describing the distribution of innovations over the space. Define H as the inverse to h: $H = h^{-1}$. Pick a given proportional imitation cost f_m then it is sufficient to prove the result focusing on a single slice of innovation space at f_m. That is, if we can show that just looking at innovations with imitation cost f_m that the welfare ratio is higher than the innovation ratio then the result must hold when looking at the whole space of innovations.

Define $I^X(f_m), W^X(f_m)$ to be the innovation and welfare levels under the regime $X = R, ZI$ when restricting to innovations with imitation cost f_m. So considering the innovation ratio we have:

$$\begin{aligned}
\text{Innovation Ratio at } f_m &= \frac{I^R(f_m)}{I^{ZI}(f_m)} \\
I^R(f_m) &= \int_0^{H(f_m)} g(f_i, f_m) df_i \\
I^{ZI}(f_m) &= \int_0^1 g(f_i, f_m) df_i
\end{aligned}$$

Turning to welfare, by the Claim above for $f_i^1 \leq H(f_m) \leq f_i^2$ we have $W^{ZI}(f_i^1, f_m) \geq W^{ZI}(H(f_m), f_m) \geq W^{ZI}(f_i^2, f_m)$. Then for some C_1, C_2 with $C_1 > 1 > C_2$ we have:

$$\begin{aligned}
W^{ZI}(f_m) &= \int_0^1 W^{ZI}(f_i, f_m) g df_i \\
&= \int_0^{H(f_m)} W^{ZI}(f_i, f_m) g df_i + \int_{H(f_m)}^1 W^{ZI}(f_i, f_m) g df_i \\
&= C_1 W(H(f_m), f_m) \int_0^{H(f_m)} g df_i + C_2 W(H(f_m), f_m) \int_{H(f_m)}^1 g df_i \\
&\leq C_1' \left(\int_0^{H(f_m)} g df_i + \int_{H(f_m)}^1 g df_i \right) \\
&= C_1' I^{ZI}(f_m)
\end{aligned}$$

Note that the inequality is strict if there are innovations both in A and outside of A, that is $\exists f_i^1 < H(f_m) < f_i^2$ with $g(f_i^j, f_m) > 0, j = 1, 2$.

Now by assumption for any $(f_i, f_m) \in A$ (i.e. with $f_i \leq H(f_m)$), $W^R(f_i, f_m) \geq W^{ZI}(f_i, f_m)$. Thus,

$$\begin{aligned} W^R(f_m) &= \int_0^{H(f_m)} W^R(f_i, f_m) g df_i \\ &\geq \int_0^{H(f_m)} W^{ZI}(f_i, f_m) g df_i \\ &= C_1' \int_0^{H(f_m)} g df_i \\ &= C_1' I^R(f_m) \end{aligned}$$

Hence we have that the Welfare ratio of R to ZI at f_m (with the inequality being strict under the condition previously stated):

$$\begin{aligned} \text{Welfare Ratio}(f_m) &= W^R(f_m)/W^{ZI}(f_m) \\ &\geq C_1' I^R(f_m)/C_1' I^{ZI}(f_m) \\ &= I^R(f_m)/I^{ZI}(f_m) \\ &= \text{Innovation Ratio}(f_m) \end{aligned}$$

□

References

A. Arundel. Patents in the Knowledge-Based Economy. *Beleidstudies Technology Economie*, 37:67–88, 2001.

James Bessen. Hold-up and Patent Licensing of Cumulative Innovations with Private Information. *Economics Letters*, 82(3):321–326, 2004.

James Bessen and Eric Maskin. Sequential Innovation, Patents, and Innovation. NajEcon Working Paper Reviews 321307000000000021, www.najecon.org, May 2006. Forthcoming in the Rand Journal of Economics.

W. Cohen, R. Nelson, and P. Walsh. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not), 2000. NBER Working Paper No. W7552.

Victor Denicolo. Two-Stage Patent Races and Patent Policy. *Rand Journal of Economics*, 31:488–501, 2000.

Giovanni Dosi. Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature*, pages 1120–1171, 1988.

Nancy Gallini. Patent Policy and Costly Imitation. *Rand Journal of Economics*, 23(1): 52–63, 1992.

Bronwyn Hall. Business Method Patents, Innovation and Policy, 5 2003.

John F. R Harter. The Propensity to Patent with Differentiated Products. *Southern Economic Journal*, 61(1):195–201, July 1994. ISSN 00384038.

Hugo A Hopenhayn and Matthew F Mitchell. Innovation Variety and Patent Breadth. *The RAND Journal of Economics*, 32(1):152–166, 2001. ISSN 07416261.

Paul Klemperer. How Broad Should the Scope of Patent Protection Be? *RAND Journal of Economics*, 21(1):113–130, 1990.

Richard Levin, A. Klevorick, R. Nelson, S. Winter, R. Gilbert, and Z. Griliches. Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity*, 3:783–831, 1987.

Edwin Mansfield. How Rapidly Does New Industrial Technology Leak Out? *Journal of Industrial Economics*, 34(2):217–223, 1985.

Peter Menell and Suzanne Scotchmer. Intellectual Property, 6 2005. forthcoming, Handbook of Law and Economics. Mitch Polinsky and Steven Shavell, eds. Amsterdam: Elsevier.

William Nordhaus. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. M.I.T. Press, 1969.

Lynne M Pepall and Daniel J Richards. Innovation, Imitation, and Social Welfare. *Southern Economic Journal*, 60(3):673–684, 1994. ISSN 00384038.

Rufus Pollock. Cumulative Innovation, Sampling and the Hold-up Problem. DRUID Working Papers 06-29, DRUID, Copenhagen Business School, Department of Industrial Economics and Strategy/Aalborg University, Department of Business Studies, 2006.

Frederic Scherer. Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation. *American Economic Review*, 62(3):422–427, 1972.

Suzanne Scotchmer and Jerry Green. Novelty and Disclosure in Patent Law. *The RAND Journal of Economics*, 21(1):131–146, 1990. ISSN 07416261.

David J. Teece. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6):285–305, December 1986.

Michael Waterson. The Economics of Product Patents. *American Economic Review*, 80 (4):860–69, September 1990.