CUMULATIVE INNOVATION, EXPERIMENTATION AND THE HOLD-UP PROBLEM

RUFUS POLLOCK

UNIVERSITY OF CAMBRIDGE

OCTOBER 6, 2008

ABSTRACT. Extending the basic model of two-stage cumulative innovation with asymmetric information to include 'experimentation' by second-stage firms, we find that the costs of a strong (versus weak) intellectual property (IP) regime may be substantially increased. In addition, these costs increase as experimentation becomes cheaper and as the differential between high and low value second-stage innovations grows, with the result that a weak IP regime is more likely to be optimal. Thus, technological change which reduces the cost of encountering and trialling new 'ideas' implies a reduction in the socially optimal level of IP rights such as patent and copyright.

Keywords: Cumulative Innovation, Hold-Up, Experimentation, Intellectual Property

JEL codes: K3, L5, O3

1. Introduction

Cumulative innovation and creativity, whereby new work build upon old, is a pervasive phenomenon. Early work on the subject by Green and Scotchmer (1995); Chang (1995) introduced a two-stage innovation model in which the second innovation is enabled by, or builds upon, the first. A central feature of these models, as well as subsequent work that extended it (such as Scotchmer (1996)), was an assumption that knowledge of costs and returns, whether deterministic or stochastic, was shared equally by innovators at different stages (i.e. was common knowledge). With common knowledge all mutually beneficial transactions are concluded, using ex ante licenses where necessary to avoid the possibility of hold-up of second-stage innovators.

This assumption, however, is problematic. If all innovators share the same information why have different innovators at different stages and why be concerned with licenses and bargaining if a single innovator could just as easily carry out all research? The simple answer is that this assumption of common knowledge does not reflect reality: many different firms engage in innovation precisely because they have specialized skills and knowledge that make it effective for them rather than another firm to engage in a given area.[1]

Here, therefore we follow the path of some recent authors in investigating cumulative innovation under asymmetric information. Specifically, we take as a starting point a 'basic' model very similar to that presented by Bessen (2004).[2] Second-stage firms are of two types (high and low value) with the type unobserved by first-stage innovators. With (strong) IP first-stage firms may require second-stage innovators to pay a royalty while with (weak) IP second-stage firms may produce without having to license from first-stage firms. As first-stage firms do not know the type of given second-stage innovator with (strong) IP there may be 'licensing failure' (that is the royalty may be set above the level that a second-stage firm is willing to pay).[3] Thus, there is a trade-off: with IP more

---

[1]See e.g. Eisenberg and Heller (1998); Hall and Ziedonis (2001); Cockburn (2005).

[2]We differ from Bessen slightly in that his focus is primarily on whether ex ante or ex post licensing occurs. Central to this analysis is his introduction of ex post royalty shares which are the royalty shares that take place in the absence of licensing. These are determined exogenously – perhaps as a policy variable or determined by invent-around costs and other factors – and Bessen shows that the socially optimal ex post royalty share is less than that obtained in ex ante bargaining (and so all licensing should occur ex post). By contrast, in our model we do not have the concept of an ex post royalty share: either a second-stage innovator obtains a license or she does not (and so then cannot produce).

[3]We note that Bessen uses the term 'holdup' to denote what we term 'licensing failure'. Since he is considering ex-ante licensing his use of the term 'hold-up' differs somewhat from the traditional usage as there are no sunk relationship-specific investments (a binding contract is possible ex-ante). Rather the 'hold-up'

first-stage innovation takes place due to the extra royalty income received by first-stage firms but some second-stage innovation may be lost as a result of 'licensing failure' due to high royalty rates.

Such a trade-off is already familiar in the literature and our main reason for presenting it is to provide a benchmark and basis for the more complex 'experimentation' model presented in the second section. 'Experimentation' is used here broadly to cover any kind of trialling and experimentation activity that is likely to take place before a license can be agreed. The logic here is that there are transactional costs and complexities involved in negotiating and executing a license that mean that it only takes place once some degree of 'experimentation' has taken place. Furthermore, experimentation benefits an innovator by increasing the probability of having a high value innovation but it is costly.

For example, a first-stage innovation might be a 'tool' which the second-stage innovator wishes to use in some manner but is unsure what the most beneficial use for this tool is or how valuable usage of the tools will be (this would be particularly relevant to Biotechnology where the issue of research 'tool' licensing is particularly prominent). 'Experimentation' in this case would correspond to trialling and investigation by the second-stage innovator in order to determine the best way to use the tool and/or how valuable such usage is. The more 'experimentation' a second-stage innovator does, the more likely the resulting use is a high value one.

Alternatively, one could imagine a first-stage innovation is a basic product that the second-stage innovator wishes to extend. Here again, the second-stage innovator while knowing that she wishes to extend a particular first-stage innovation may not necessarily have a clear idea as to how best to do this (or whether the particular idea she has is actually feasible). 'Experimentation' would then indicate the trialling and investigation necessary to reduce these uncertainties and improve the likelihood the result is a good one.

Real-world examples of such situations abound. Consider, for example, software where it is common for developers to expend significant time trying out and experimenting

---

is simply that, just like a monopolist facing heterogeneous consumers, a first-stage innovator is facing a set of second-stage innovators with private and heterogeneous values and so may set a profit-maximizing royalty rate that excludes some second-stage innovators from licensing. Since, the 'experimentation' case we discuss below resembles more closely a traditional 'hold-up' situation we prefer to reserve that term for use there and to to use 'licensing failure' for the situation described here.

with an existing product or 'library' in order to determine whether they can extend it or integrate in the way that would be useful or fits with their existing needs. Significant effort may have to be expended before a formal license is concluded (for example, the license may depend on the exact intended usage), and the more time a user or developer spends experimenting the more likely the resulting application is a good one. The same logic applies to other information products such as databases, as well as to research tools in areas like the life sciences. Here a (potential) user may need to spend significant time exploring the content and features of the product, as well as trialling different ways to use and apply it prior to agreeing a formal license, with the likelihood of a good application increasing in the time spent in this way.[4]

Returning to the model, the crucial point is that experimentation takes place prior to any kind of royalty negotiation. As a result, strong IP rights which require second-stage innovators to license may now have cost in addition to that arising from licensing failure: fearing high royalty rates second-stage firms will reduce the level of experimentation they do and thereby reduce the average quality of second-stage innovations. We should emphasize here that this is not a simple hold-up problem – while royalty setting occurs after experimentation it still takes place before the primary investment decision by second-stage innovators – but derives from an *interaction* of 'licensing failure' with the 'experiment' effect on innovation quality. Importantly, since reduced experimentation affects all second-stage innovators, the consequences for welfare may be substantially greater those of traditional 'licensing failure'. In this sense, experimentation acts as a 'multiplier' of licensing failure, greatly increasing the potential losses from strong IP.

We also find some interesting comparative statics. For example, the lower the experimentation costs or the larger the differential between high and low value second-stage innovations, the more likely it is that a regime with weak or no intellectual property rights will be preferable. Thus, in the context of this model, technological change which reduces

---

[4]Consider here the case of geographic data in the UK, where the primary source for such data is the Ordnance Survey. The Ordnance Survey offer a range of possible reuse licenses, termed Specific Use Contracts, each targeted at a particular area. In all there are 14 options with, for example, one for Navigation Products, one for Location Based Services, one for User Derived Datasets etc. The terms and conditions vary between the contracts and it is therefore important for a reuser to know which contract to choose. However at the early stages of the development of a new product or application it may costly (or simply impossible) to know exactly which contract should be used and it will only be later when much of this uncertainty has been resolved that it will be worth incurring the cost and effort of negotiation to agree with the Ordnance Survey which SUC to use.

the cost of encountering and trialling new 'ideas' should imply a reduction in the socially optimal level of intellectual property rights such as patents and copyright.

2. Related Literature

In addition to the works already mentioned our paper has connections with some other parts of the innovation literature. For example, our comparison of strong and weak IP regimes has analogies with the infringing/non-infringing comparison of Denicolo (2000).

A stronger connection exists to Bessen and Maskin (2006). They investigate the welfare impact of 'licensing failure' due to asymmetric information in a model of cumulative innovation and show that, with cumulative innovation, in contrast to what occurs in a 'one-shot' model, IP may, in some circumstances, reduce rather than increase innovation (and social welfare). However their focus is rather different from ours (complementarities in research rather than experimentation) and their results arise for different reasons. Specifically, in their model there are multiple stages with (the same) two firms at each stage. Each may choose to participate or not in researching the current innovation and the next innovation stage is reached if, and only if, research at the current stage is successful, with success an increasing function of the number of participating firms. As a result their is an 'externality' from participation in a given stage: though the value of success at the current stage accrues only to the winning firm by enabling subsequent stages (some of which may be won by the other firm) success also increases the other firms expected revenue. As a result, when one firm is excluded from subsequent stages due to 'licensing failure' under an IP regime the effect on welfare can be far more severe than in the one-stage case.

The present paper also has a relationship to the recent work of Polanski (2007). That focused on comparing 'Open-Source' (OS) and 'Proprietary' (PR) production. While some of the results here have a similar flavour to those found there, particularly in the way that the negative effects of IP can 'multiply up'. However, our model differs substantially both in form and substance. In particular, here, in constrast to Polanski (2007), there are only two stages and the main 'hold-up' issue arises in relation to the second-stage innovators 'experimentation' effort and its interaction with licensing failure in the presence of imperfect information about the types of second-stage innovators.

Finally, we should point out that our results relate to a variety of recent policy debates. For example, in December 2006 the Gowers Review of Intellectual Property which had been setup by the UK government to examine the UK's current IP regime, provided, as one its recommendations (no. 11), that "Directive 2001/29/EC [the EU Copyright 'InfoSoc' Directive] be amended to allow for an exception for creative, transformative or derivative works, within the parameters of the Berne Three Step Test." Such a 'transformative use exception' would correspond very closely to the weak IP regime considered in the model presented here. Meanwhile in 2005 in the United States, the Supreme Court in *Merck KGaA v. Integra Life Sciences I, Ltd*[5] created a very broad research exemption in relation to pre-clinical R&D. Such a change again corresponds closely in the model to a move towards a weak IP regime in which a second-stage product would not infringe on a first-stage firm's patent.

3. A Basic Model of Two-Stage Cumulative Innovation

3.1. **The Model.** We adopt a simple model of two stage innovation in which the second innovation builds upon the first in some manner – either as an application or as an extension of it. All agents are risk-neutral and act to maximize profits.

Innovations are described by their net (stand-alone) value v (revenue minus costs). Because our interest lies in examining the trade-off between innovation at different stages we make no distinction between social and private value (i.e. there are no deadweight losses) and v may be taken to be both.

We assume the base (first) innovation has two possible (stand-alone) values: low (v_1^L) and high (v_1^H) with probability $p, (1 - p)$ respectively. We assume that $v_1^L < 0$ so that without some additional source of revenue, for example from licensing (see below), the innovation will not be produced. High value innovations have positive stand-alone value, $v_1^H > 0$, and so do not require an outside source of revenue in order to be profitable.

Second-stage innovations also take two values: low (v_2^L) and high (v_2^H) with probability $q, (1 - q)$ respectively and $v_2^H > v_2^L > 0$. While the value of a second-stage innovation is known to the innovator who thinks of it, the value is **not** known to the owner of the first-stage innovation which it builds upon (this could occur because of imperfect information regarding revenue, costs or both). Without loss of generality we shall assume that the

---

[5] The full opinion is available at http://www.supremecourtus.gov/opinions/04slipopinion.html.

number (or measure) of second-stage innovations per first-stage innovation is one (having N second-stage innovations per first-stage innovation would just require replacing v_2^H with Nv_2^H and v_2^L with Nv_2^L). We also assume that $v_1^L + v_2^L \geq 0$ – this ensures that whatever the value of q the overall value generated by a first-stage innovation is positive (the overall value is the stand-alone plus the value of dependent second-stage innovations).[6]

3.1.1. *Intellectual Property Rights and Licensing.* We wish to consider two regimes: one with (strong) intellectual property rights (IP) and one with weak or no intellectual property rights (WIP). With intellectual property rights every second-stage innovator will require a license from the relevant first-stage innovator in order to market her product, while without intellectual property rights she may market freely without payment or licence.[7]

We assume that the stand-alone value of the first and second-stage innovators (v_i) are unaffected by the intellectual property rights regime. Note here, that strictly, we only require that the *sum* of the first-stage and (dependent) second-stage innovation values is unaffected by the IP regime as we can incorporate transfers via changes in the probability p that a first-stage innovation is low value (under strong IP any gains by second-stage firms at the expense of first stage firms will be taken back through the royalty so p only plays a role in the weak IP regime). Thus, simple business stealing, in which the total combined rents of the two stages remain unchanged, but where the second-stage takes rents from the first-stage, is permitted in this model. However, we do not allow for the general erosion of overall rents under the weak IP regime due to a general increase in direct product market competition.[8] It is therefore best to take the weak IP regime as corresponding to non-infringement where distinct improvements or applications do not need to license but where there is still of protection against direct copying.

Finally, we take the licence to define a lump-sum royalty payment r. This assumption is without loss of generality since, in this model, an innovation is entirely defined by its

---

[6]Allowing values of v_1^L less than v_2^L does not alter the analysis in any significant way but brings extra complexity to the statement and proof of propositions.

[7]Given that we are dealing with cumulative innovation some readers might prefer the infringing (I) vs. non-infringing (NI) dichotomy with its implication of a distinction between 'horizontal' imitation and 'vertical' improvement or application of a product.

[8]This is similar to the approach taken explicitly or implicitly in much of the previous literature, for example Green and Scotchmer (1995); Chang (1995); Scotchmer (1996); Denicolo (2000); Bessen (2004) We would also note that with rent dissipation one must immediately provide for a distinction between private rent and social value further complicating the analysis, and distracting from the focus of this paper which is on the welfare effects of IP purely as they operate through the impact on innovation leaving aside the (classic and well-known) direct effect of IP on welfare due to the deadweight loss of reduced access.

Player	Second-Stage Innovator					
	Value Type			Low (q)		High (1-q)
First Stage		Action	NI	I	NI	I
Innovator	Low (p)	r	v_1^L , 0	$v_1^L + r$, $v_2^L - r$	v_1^L , 0	$v_1^L + r$, $v_2^H - r$
	High (1-p)		v_1^H , 0	$v_1^H + r$, $v_2^L - r$	v_1^H , 0	$v_1^H + r$, $v_2^H - r$

TABLE 1. Action and Payoff Matrix Assuming First-Stage Innovator Invests. (I/NI = Invest/Do Not Invest, r = Royalty Rate)

net value v and there are no other attributes available to use in designing a mechanism to discriminate between types of second-stage innovator.[9] The royalty is set ex-ante, that is prior to the second-stage innovator's decision to invest, and is in the form of a take-it-or-leave it offer by the first-stage innovator.

3.1.2. *Sequence of Actions.* The sequence of actions in the model is:[10]

  (1) Nature determines the value type of the first-stage innovator.

  (2) A first-stage innovator decide whether to invest. If the first-stage innovator does not invest the game ends and all payoffs are zero. Assuming the first-stage innovator invests the game continues.

  (3) The first-stage innovator sets the royalty rate r (under the no/weak IP regime second-stage innovations do not infringe and so the de facto royalty rate is 0).

  (4) Nature determines the value type of a second-stage innovator (unobserved by first-stage).

  (5) Given the royalty rate and their type second-stage firms decide whether to invest.

  (6) Payoffs are realized.

The action/payoff matrix is summarized in Table 1.

3.2. **Solving the Model.** Define a constant, α, as follows:

$$\alpha \equiv \frac{v_2^H - v_2^L}{v_2^H}$$

Proposition 3.1. *With intellectual property rights, the game defined above has the following Subgame Perfect Nash equilibria. A second-stage innovator invests if and only if its*

---

[9]For example, there are no quantities on which to base a non-linear pricing scheme (fixed fee plus per unit fee royalty). For the same reason there is no opportunity to use type-contingent menus, or any other form of more complex licensing agreement, to increase total royalty income by discriminating between high and low value innovators.

[10]Stages 3 and 4 may be swapped without in any way affecting the results (this ordering may be more appealing to some readers).

	RL	**RH**
IP	$v_1 + v_2$	$v_1 + (1-q)(v_2^H)$
WIP	$(1-p)(v_1^H + v_2)$	$(1-p)(v_1^H + v_2)$
IP - WIP	$p(v_1^L + v_2^L) + p(v_2 - r_L) \geq 0$	$p(v_1^L + (1-q)v_2^H) - (1-p)qv_2^L$

TABLE 2. Welfare in the Basic Model

realized value is greater than or equal to the royalty rate (i.e. net profits are non-negative).
A first-stage innovator invests and sets a low royalty rate (RL), $r_L = v_2^L$ if the probability
of a low value innovation (q) is greater than α and a high royalty rate (RH) $r_H = v_2^H$ if
$q \leq \alpha$. When $q = \alpha$ the first-stage innovator may set any royalty of the form r_L with
probability x and r_H with probability $1 - x$, $x \in [0,1]$. Thus, there always exist a pure
strategy equilibrium and except when $q = \alpha$, this equilibrium is unique.

Proof. See appendix. □

Proposition 3.2. *Without intellectual property rights the game above has the following*
solution: both types of second-stage innovators invest but, of first-stage innovators, only
those that have 'high-value' innovations invest (there are $1 - p$ of these type).

Proof. Trivial. □

3.3. **Welfare.** To determine welfare we need to know the 'trade-off' between first and
second-stage innovations that occurs when revenue is allocated from one to the other
by licensing. As stated above, without royalty income from second-stage innovations a
proportion p of first-stage innovations are not produced with average (stand-alone) value
v_1^L. The remaining innovations $(1-p)$ are produced irrespective of whether royalty revenue
is received and have average value v_1^H.

Let us now consider social welfare in the four possible situations given by (IP, RL), (IP,
RH), (WIP, RL), (WIP,RH) as well as the difference in welfare between an intellectual
property regime and a weak intellectual property regime (IP-WIP). Due to our earlier
assumption welfare is determined by calculating total net value. Define for convenience
$v_1 = pv_1^L + (1-p)v_1^H$, the average first-stage innovator value (if all innovate), and $v_2 = qv_2^L + (1-q)v_2^H$, the average second-stage innovator value (if all innovate). We summarize
the welfare situation in Table 2.

3.4. **Policy Implications.**

Proposition 3.3. *When a low royalty will be set ($q \geq \alpha$) a strong IP regime is optimal.*

Proof. In the low royalty (RL) situation all second-stage innovations will be produced whether there is IP or not. In that case one wishes to maximize returns to the first innovator and patents do this by transferring rents via licensing. Formally in the low royalty case the welfare difference between patents and no patents (IP-WIP) is:

$$p(v_1^L + r_L) + p(v_2 - r_L)$$

Both of the terms in brackets are positive implying that the intellectual property regime delivers higher welfare than the weak intellectual property (WIP) regime.   □

The situation when the high royalty will be set is less clear. Define β as the proportion of the royalty payment to a *low-value* first-stage innovator that would be 'used up' in paying their extra costs:

$$\beta \equiv \frac{-v_1^L}{(1-q)r_H}$$

Note that v_1^L is negative and must be less in absolute terms than the royalty received $(1-q)r_H$ as we are assuming that the royalty enables low value first-stage innovators to produce. Under this definition $\beta = 1$ corresponds to the case where all of the royalty paid to a low-cost first-stage innovator being used to pay their 'extra' costs while $\beta \approx 0$ means all of the royalty payment is being retained as extra profits (and welfare). With this definition the surplus from each extra first stage innovation made possible by strong IP in the RH case is $v_1^L + (1-q)v_2^H = (1-\beta)v_2^H$.

Proposition 3.4. *When a high royalty will be set ($q < \alpha$) an intellectual property regime will be preferable to a weak intellectual property (WIP) regime if and only if (NB: in fact with equality one would be indifferent):*

$$p \geq \frac{qv_2^L}{qv_2^L + (1-\beta)(1-q)v_2^H} \tag{3.1}$$

$$= \frac{Licensing\ Failure\ Cost}{Licensing\ Failure\ Cost + Surplus\ From\ Extra\ 1st\ Stage} \tag{3.2}$$
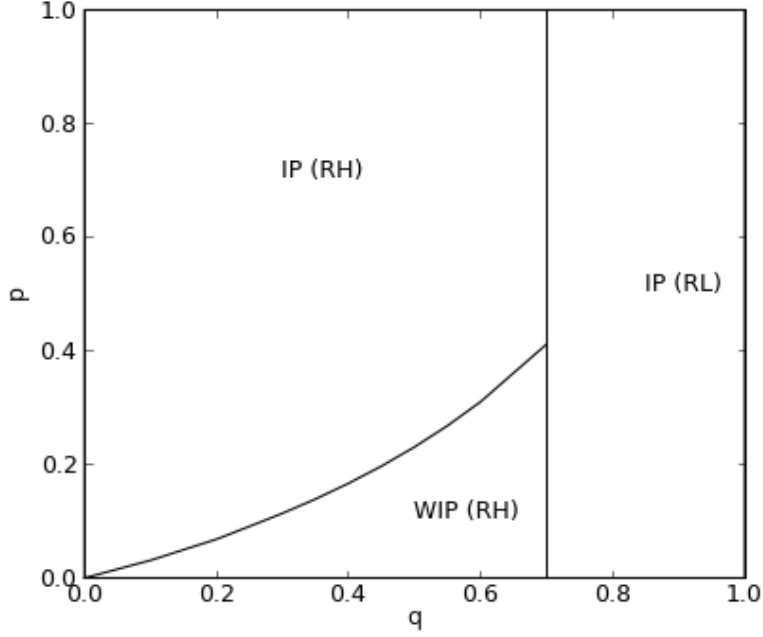
FIGURE 1. Optimal policy as a function of the probabilities of low value first-stage (p) and second-stage (q) innovations. α equals 0.7 so to the right of the line $q = 0.7$ a low royalty will be set (RL) and a strong IP regime is optimal. To the left of that, the region is divided into two parts by the line defined by the equation in Prop 3.4 (with $\beta = 0$).

Proof. From Table 2 a strong IP regime yields higher welfare than an WIP regime if and only if:

$$p(v_1^L + (1 - q)v_2^H) \geq (1 - p)qv_2^L$$

Making p the subject of this inequality and using β we obtain the stated result. $\square$

We represent the import of these propositions graphically in Figure 1, a diagram which shows optimal policy regions as a function of the exogenous probabilities of low value first-stage (p) and second-stage (q) innovations. As this shows, for the standard case corresponding to low β, for most values of p and q, the (strong) IP regime is preferable to the weak or no IP regime. Putting this the other way round, it is only under very particular circumstances that a weak or no IP regime is preferable: most first stage innovation must occur anyway – p low – *and* the proportion of second-stage innovations that are of low value (and hence suffer licensing failure) must be substantial but not too great – q large but not bigger than α.

4. Cumulative Innovation with Experimentation

The 'experimentation' model differs from the 'basic' one presented in the previous section by the addition of experimentation by second-stage firms. As discussed in the introduction, this experimentation takes place at an early stage, in particular, prior to any royalty setting, and it affects the quality of the innovative idea a second-stage firm has. However, as in the 'basic' model all investment required to realize the innovation as an actual product, and all revenue generated from doing so, occurs after royalty setting.[11] Formally, the new sequence of actions is a follows (modifications compared to the basic model are bolded for clarity):[12]

(1) Nature determines the value type of the first-stage innovator.

(2) A first-stage innovator decide whether to invest. If the first-stage innovator does not invest the game ends and all payoffs are zero. Assuming the first-stage innovator invests the game continues.

(3) **Second-stage innovators chooses their level of experimentation k.**

- Experimentation has constant marginal cost τ.

- With regard to knowledge of the experimentation level chosen by a second-stage firm there are two possibilities: either 1) first-stage innovator observe the experimentation level or 2) they do not. In what follows we focus on the case where the experimentation level is unobserved as we feel this is more realistic though the results are unchanged (and simpler to derive) when it is observed.

- Experimentation affects the probability, q, that a second-stage firm turns out to have a low value innovation: $q \equiv q(k)$.

---

[11]We have also examined the case where the royalty is set prior to experimentation. Clearly, this takes one back to a setup very close to the basic model in that all royalty setting is ex-ante. Unsurprisingly, therefore, the effect on experimentation is less than in the case examined here since a first-stage innovator takes account (at least to some extent) of the impact of their actions on experimentation when setting the royalty. However, as argued in the introduction there are many cases where it is likely that at least some costly activity ('experimentation') takes place prior to any royalty negotiation and it is therefore under that assumption that we proceed. Furthermore, the case where the royalty is set completely ex-ante (i.e. before any kind of costs are incurred) can be incorporated into the present framework simply by assuming that the 'experimentation' as modelled here is unimportant, i.e. in the terminology below: $q(k)$ varies very little with k.

[12]We remark that as for the basic model, reversing steps 4 and 5 – which may, for some readers, be the more natural ordering – makes no difference to the solution of the model.

Player			Second-Stage Innovator			
			Sample (k)			
	Value Type		Low (q(k))		High (1-q(k))	
First-Stage		Action	NI	I	NI	I
Innovator	Low (p)	r	v_1^L , $-k\tau$	$v_1^L + r$, $v_2^L - r - k\tau$	v_1^L , $-k\tau$	$v_1^L + r$, $v_2^H - r - k\tau$
	High (1-p)		v_1^H , $-k\tau$	$v_1^H + r$, $v_2^L - r - k\tau$	v_1^H , $-k\tau$	$v_1^H + r$, $v_2^H - r - k\tau$

TABLE 3. Action and Payoff Matrix Assuming First-Stage Innovator Invests (I/NI = Invest/Do Not Invest, r = Royalty Rate)

- Properties of $q(k)$: $q' \leq 0$ (otherwise no benefit from experimentation) and there are diminishing returns to experimentation: $q'' \geq 0$. If no experimentation takes place all firms are of low value type ($q(0) = 1$).[13] The functional form $q(k)$ is assumed to be common knowledge.

(4) The first-stage innovator sets the royalty rate r (under the weak IP regime second-stage innovations do not infringe and so the de facto royalty rate is 0).

(5) Nature determines the value type of a second-stage innovator (high or low value: v_2^H, v_2^L). As before, nature's move determines a firm's type with certainty. However, here the ex-ante probability of being low-value (q) is not exogenously given (as in the basic model) but determined endogenously by the experimentation level.

(6) Given the royalty rate and their type second-stage firms decide whether to invest.

(7) Payoffs are realized.

The new action/payoff matrix is shown in Table 3. The rest of the section is organized as follows. In the first part we derive the equilibrium of the game just described (main results in Propositions 4.1 and 4.2). Following this we analyse welfare and, as in the basic model, relate the optimal regime (strong IP or weak IP) to a simple expression in the main exogenous variables. Finally we end the section by exploring how this expression, and hence the likelihood that strong/weak IP is optimal, varies with experimentation costs and the distribution of innovation values.

4.1. **Equilibrium.** Define, as in the basic model, a high royalty to be equal to the value of a high-value second-stage innovation: $r_H = v_2^H$, and a low royalty to be equal to the value of a low-value second-stage innovation: $r_L = v_2^L$. Recall also from the basic model the variable α, the value of q – itself the probability a second-stage innovation is low-value

---

[13]It should be noted that this assumption is not important for the results and has simply been made for simplicity. It is easy to adapt all of the results to the case where even under zero experimentation there is a non-zero probability of having a high-value innovation.

– at which a first-stage firm is indifferent between setting a high and a low royalty rate:

$$\alpha \equiv \frac{v_2^H - v_2^L}{v_2^H}$$

Relatedly, here we define k_α as the experimentation level such that the probability a second-stage innovation is low-value is α, i.e. $q(k_\alpha) = \alpha$. Lastly define the experimentation level k_2 implicitly by:

$$q'(k_2) = \frac{-\tau}{v_2^H - v_2^L}$$

k_2 is the level of experimentation undertaken in the case where both types of second-stage innovators (high and low-value) will find it worthwhile to invest and market their innovation/product (for formal derivation of this see results in appendix). It therefore occurs, as we shall see, either in the case when the royalty is sufficiently low (which obviously include the case of weak IP where the royalty is zero). It is also, therefore, the level of experimentation which maximizes expected second-stage innovation value. As such, is the socially optimal experimentation level. Lastly, we observe that both k_α and k_2 are defined in terms of exogenous variables.

Proposition 4.1 (Equilibrium under weak IP). *Under weak IP the 'experimentation' model has the following solution: second-stage innovators sample at level k_2 and both types of second-stage innovators invest. Of first-stage innovators, those that have 'high-value' innovations invest (there are $1 - p$ of these type) and those with 'low-value' innovations do not.*

Proof. Follows trivially from best-response correspondences derived in Proposition B.3 (see appendix). □

This intuition behind this result is straightforward. With no royalty needing to be paid, at the second-stage both high and low type innovators will invest, and, knowing this, they will experiment at level k_2 (this was how k_2 was defined). Lastly, just as in the basic model, without royalty income only the $1 - p$ high (stand-alone) value first-stage innovators invest.

Proposition 4.2 (Equilibrium under strong IP). *With intellectual property rights (IP) the perfect Bayesian equilibrium of the game defined above falls into one of two cases:*

(i) **Low royalty case ($k_2 \leq k_\alpha$)**

  (1) *First-stage innovators: both high and low value types invest, believe that second-stage innovators sample at level k_2 and set a low royalty rate.*

  (2) *Second-stage innovators: sample at level k_2 and both high and low value types invest.*

(ii) **Mixed royalty case ($k_2 > k_\alpha$)**

  (1) *First-stage innovators: both high and low value types invest, believe that second-stage innovators sample at level k_α and set a mixed royalty rate consisting of a high royalty (r_H) with probability x_α and a low royalty (r_L) with probability $(1 - x_\alpha)$ where:[14]*

$$x_\alpha = 1 - \frac{\tau}{-q'(k_\alpha)(v_2^H - v_2^L)}$$

  (2) *Second-stage innovators: sample at level k_α and invest if and only if the realized value of their innovation is greater than the royalty rate (though the first-stage innovator is playing a mixed strategy the second-stage innovator knows the royalty rate with certainty at the point of investment).*

Proof. See appendix.   □

What is the intuition behind this result. Assume first that second-stage firms have a pure experimentation strategy and that first-stage firms are restricted to setting a mixture of a high or low royalty (these facts are proved rigorously in the appendix). Now, k_α here plays the role of α in the basic model. When second-stage firms experiment at a level above k_α first-stage firms will want to set a high royalty, when experimentation is below k_α they will set a low royalty, and at k_α first-stage firms will be indifferent between a high and low royalty.

Conversely, second-stage firms will experiment at level k_2 when a low royalty is set, at level 0 when a high royalty is set and at some level in between when a mixed royalty is set. Putting this together we have two best-response functions (one for royalty setting and one for experimentation). Looking for equilibria we see that: (1) if $k_2 \leq k_\alpha$ and second-stage firm experiment at k_2, first-stage firms, (who correctly expect this level of

---

[14]Note examining the definition of k_2 shows that $k_\alpha < k_2$ guarantees that x_α is non-negative.

experimentation) will set a low royalty, and this royalty, is, in turn, compatible with second-stage firms experimenting at level k_2 (2) if $k_2 > k_\alpha$ the only possible way to have an equilibrium is with first-stage firms setting a mixed royalty, but this is possible only if second-stage firms experiment at level k_α. This in turn defines the mixed royalty strategy, x_α, by the need to have second-stage firms sampling at k_α. Thus, in this second case, the possibility of future licensing failure when a high royalty is set reduces the effort that second-stage firms are willing to invest in experimentation at the initial stage (from k_2 to k_α). As we shall see this has significant implications for welfare.

4.2. **Welfare.** For the welfare calculations we proceed as in the original model. Specifically, recall that the exogenous variable p gives the proportion of first-stage innovations which require the royalty income from strong IP in order to be viable (they are of low stand-alone value). We will focus on obtaining the 'marginal' or 'break-even' level of p, denoted p^m. This is the value of p that separates the two regimes, i.e. for p above p^m a strong IP regime is optimal and for lower p a weak IP regime is optimal.

As for the basic model define $v_1 = pv_1^L + (1-p)v_1^H$ and $v_2(k) = -k\tau + (1-q(k))v_2^H + q(k)v_2^L$. These are the expected (stand-alone) value of a first-stage innovator and the expected value generated by a second-stage innovator experimenting at level k.

Proposition 4.3. *[The Optimal Regime in the Low Royalty Case] In the low royalty case ($k_2 < k_\alpha$) $p^m = 0$ and it is always (weakly) optimal to have a strong IP regime.*

Proof. See appendix.   □

This intuition here is simple: when a low royalty is set under strong IP there is no licensing failure and second-stage firms experiment at the optimal rate. Thus, the strong IP situation is at least as good as the weak IP case and will be better whenever some first-stage firms need royalties in order to innovate ($p > 0$). The mixed royalty case is more complex – and more interesting. The key result is given by the following proposition.

Proposition 4.4. *[The Optimal Regime in the Mixed Royalty Case] In the mixed royalty case ($k_2 \geq k_\alpha$) the break-even proportion of first-stage innovations requiring royalty income*

p^m *is given by:*

$$p^m \equiv \frac{(v_2(k_2) - v_2(k_\alpha)) + x_\alpha q(k_\alpha)v_2^L}{(v_2(k_2) - v_2(k_\alpha)) + x_\alpha q(k_\alpha)v_2^L + (v_2(k_\alpha) - x_\alpha q(k_\alpha)v_2^L - (-v_1^L))} \tag{4.1}$$

$$= \frac{Reduced\ Exper.\ Cost + Licensing\ Failure\ Cost}{Reduced\ Exper.\ Cost + Licensing\ Failure\ Cost + Surplus\ from\ Extra\ 1st\ Stage}$$

(4.2)

And a strong (weak) IP regime is optimal if $p > p^m$ *(*$p < p^m$*).*

Proof. See appendix.   □

It is valuable to compare equation (4.1) with equation (3.1) from the basic model. The main, and most obvious, difference is that, as well as the standard 'licensing failure cost' of (strong) IP, there is another, additional, cost in the form of 'reduced experimentation' (and reduced average value of second-stage innovations). As shall become clear below this additional term makes a substantial difference to whether strong IP is preferred to weak IP. First let's discuss the terms in the equation in a little more detail.

Reduced Experimentation Cost: $v_2(k_2)$ is the average value of second-stage innovations when second-stage firms sample at the unrestricted (and optimal) level k_2. Under the strong IP regime second-stage firms sample at the restricted level k_α due to the higher (average) expected royalty. Thus, the average value of a second-stage innovation is less under the IP regime compared to the weak IP regime due to this reduced experimentation precisely by the amount: $v_2(k_2) - v_2(k_\alpha)$ (NB: obviously this only applies to those second-stage innovations associated with the $(1 - p)$ first-stage innovations which are produced under both the strong IP and the weak IP regime.)

Licensing Failure Cost: licensing failure occurs when a second-stage firm with a low-value innovation is faced with a high royalty rate and therefore exits. Under the strong IP regime, x_α is the probability that a high royalty is set by a first-stage innovator and $q(k_\alpha)$ is the probability a second-stage firm has a low-value innovation. Thus $x_\alpha q(k_\alpha)$ is the probability that licensing failure occurs. When it does the loss equals the potential value of the second-stage innovation: v_2^L.

Surplus from Extra First-Stage Innovation: the plus side of the strong IP regime is the extra first (and dependent) second-stage innovation which occurs because first-stage innovators receive royalty income. The proportion of first-stage innovators who require

royalty income to be viable is p. For each such innovation the net surplus generated equals the first-stage firm's standalone value: v_1^L, plus the surplus generated by dependent second-stage firms: $v_2(k_\alpha) - x_\alpha q(k_\alpha) v_2^L$ (the average value if all second-stage firms produced: $v_2(k_\alpha)$, minus the surplus of those second-stage firms who are held-up: $x_\alpha q(k_\alpha) v_2^L$).

4.3. **Policy Implications.** We do not have any precise estimates for the exogenous parameters such as the experimentation cost (τ) or the values of second-stage innovations (v_2^H etc). Thus, we cannot make direct statements about which regime would yield higher welfare (whether at the global or industry level). However we can look at how p^m, the marginal value of p, varies with the other exogenous parameters.

Our general results are summarized in Figure 2 and Figure 3 (the formal details are in Propositions 4.5 and 4.7 below). As we note in the captions, one can only indicate the general form as any specific curve for p^m will depend on the functional form for q and of course the values of the other exogenous parameters. Nevertheless, both diagrams have clear implications.

As detailed in Corollary 4.8, the first diagram shows that, as experimentation costs become low, it becomes more likely that a weak IP regime is preferable. In particular, as experimentation costs τ get very low, weak IP is optimal in all industries except those where p, the proportion of first-stage innovations *dependent* on royalty income, is very high. Such cases of high p would correspond to a situation where the initial (first-stage) innovations have little standalone value but generate revenue as tools or components for follow-up innovations. Conversely, in industries where a reasonable number of first-stage product have substantial standalone value – for example, the music industry – it will be optimal to have a weak IP regime in which second-stage innovators are free to reuse first-stage innovations without needing to license.

Coming to the second figure, as detailed in Corollary 4.9 this shows that, as high-value second-stage innovations become more important relative to low value ones (the distribution is more skewed of innovation values are more skewed), it is more likely that a weak IP regime is socially optimal. Specifically, unless the proportion of first-stage
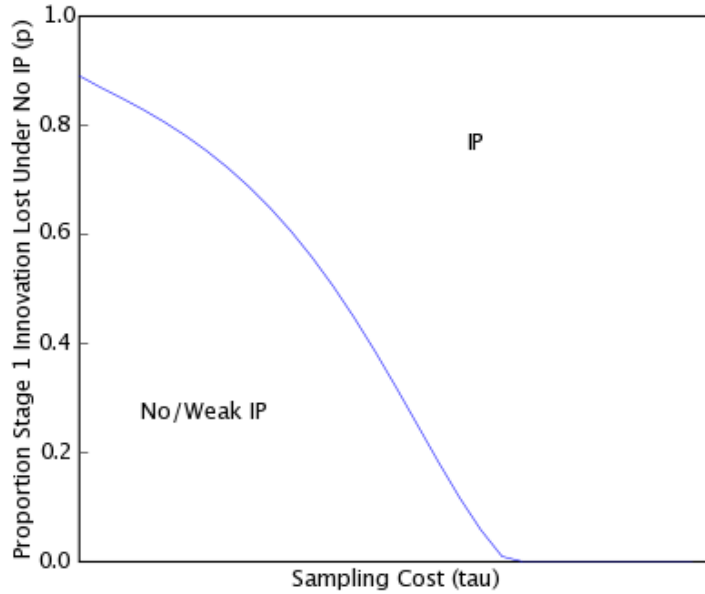
FIGURE 2. Marginal p^m as a function of experimentation cost. Ticks on τ axis have been specifically omitted as they would be misleading – any particular value will depend on parametrization, functional form for q etc. However in this specific case we note that $v_2^H = 10, v_2^L = 1.0, q(k) = e^{-k}$ and $p^m = 0$ above 8.0.

innovators *dependent* on income from second-stage firms is very high (approximately 100% for very skewed value distributions) it will be optimal to have a weak IP regime.[15]

The contrast here with the results of the 'basic model', as represented in Figure 1), is sharp: in the 'basic model' for all but very low values of p (when all first-stage innovations happen anyway) a strong IP regime is preferable. Here, at least when experimentation costs are low or the value of innovations is highly skewed, a weak IP regime is preferable for all but very high values of p (which corresponds to the case when almost no first-stage innovations would occur without licensing income). Thus, the introduction of experimentation into the two-stage model substantially alters the trade-off between (strong) IP and weak (or no) IP, and making it much more likely that a weak (or no) IP regime is preferable.

---

[15]It is noteworthy here that many studies of the value of intellectual property rights indicate that their distribution is highly skewed with a few very high value works and many low value works, see Pakes (1986); Schankerman and Pakes (1986); Schankerman (1998). This suggests that values are indeed highly skewed, i.e. $v_2^H >> v_2^L$.
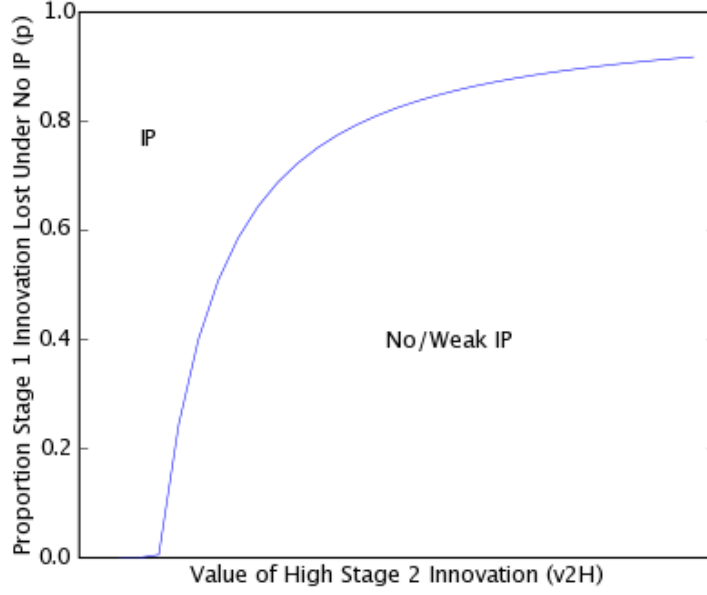
FIGURE 3. Marginal p^m as a function of v_2^H (or equivalently, for fixed v_2^L: $v_2^L - v_2^L$). For the same reasons given in relation to Figure 2 ticks on the v_2^H axis have been omitted. However to give the reader some sense of proportion we note that $\tau = 0.5, v_2^L = 1.0, q(k) = e^{-k}$ and $p^m = 0$ below approximately 3.0.

We now turn to the task of deriving formally the results shown in the two figures via the following propositions.

Proposition 4.5. *The experimentation levels k_α and k_2 have the following comparative statics:*

$$\frac{dk_\alpha}{d\tau} = 0, \frac{dk_\alpha}{dv_2^H} < 0$$

$$\frac{dk_2}{d\tau} < 0, \frac{dk_2}{dv_2^H} > 0$$

And taking limits:

$$\lim_{\tau \to \infty} k_2 = 0, \lim_{v_2^H \to v_2^{L+}} k_2 = 0, \lim_{\tau \to 0} k_2 = \infty, \lim_{v_2^H \to \infty} k_2 = \infty \qquad (4.3)$$

$$\lim_{v_2^H \to \infty} k_\alpha = 0 \qquad (4.4)$$

Proof. Recall that we have:

$$q(k_\alpha) = \frac{v_2^H - v_2^L}{v_2^H}, q'(k_2) = \frac{-\tau}{v_2^H - v_2^L}$$

Given that $q' < 0$ and $q'' > 0$ the results following trivially by simple differentiation. $\square$

Remark 4.6. The intuition behind these results is straightforward. k_α is the level of experimentation that leaves a first-stage innovator indifferent between charging a high and a low royalty rate. As such it is a function only of the relative values of the two types of innovation (and of q) and does not depend on the cost of experimentation at all.

The intuition in the second case is a little more complicated. If we increase v_2^H keeping v_2^L constant we increase the differential between high and low value second-stage innovations. Then the net change in revenue for a first-stage innovator's from switching to a high royalty rate must increase (loss of royalty revenue from low-value second-stage innovations is lower relative to royalty from high-value second-stage innovations). Hence, the proportion of high value second-stage innovations $(1 - q(k))$ at which the switch to a high royalty rate is made is smaller and the corresponding level of experimentation (k_α) is smaller.

Coming to k_2, which is the optimal level of experimentation (and that performed under a low or zero royalty), we have unsurprisingly that as the cost of experimentation goes down the amount of experimentation goes up. Similarly, an increase in the relative size of a high value innovation compared to a low value one, increases the benefit of experimentation and therefore increases the amount of experimentation performed.

Combining the differentials with the limits we have that (a) keeping other variables fixed there exists a unique finite τ^* such that for $\tau < \tau^*, k_2 > k_\alpha$ and a mixed royalty is set (conversely for $\tau > \tau^*$ a low royalty is set and $p^m = 0$); (b) similarly there exists a unique v^* such that for $v_2^H > v^*, k_2 > k_\alpha$ and a mixed royalty is set (conversely for $v_2^H < v_2^{H*}$ a low royalty is set and $p^m = 0$). This then demonstrates the validity of the right-hand part of Figure 2 and the left-hand part of Figure 3 where we have $p^m = 0$.

What occurs then if $k_2 > k_\alpha$ and we are in the mixed royalty case?

Proposition 4.7. *Assuming $k_2 > k_\alpha$ (i.e. τ sufficiently small or v_2^H sufficiently large) then:*

$$p \geq p^m \equiv \frac{(v_2(k_2) - v_2(k_\alpha)) + x_\alpha q(k_\alpha) v_2^L}{(v_2(k_2) - v_2(k_\alpha)) + x_\alpha q(k_\alpha) v_2^L + ((v_2(k_\alpha) - x_\alpha q(k_\alpha) v_2^L) - (-v_1^L))}$$

And we have that:

$$\frac{dp^m}{d\tau} < 0$$
$$\frac{dp^m}{dv_2^H} > 0$$

That is, the marginal level of first-stage innovation lost under weak IP (that is the level such that above this a strong IP regime is optimal) is (a) decreasing in experimentation costs (b) increasing in the relative size of high value to low value second-stage innovations.

Proof. See appendix. □

Informally this result can be explained as follows. Reductions in experimentation costs will increase the 'optimal' level of experimentation (k_2) relative to the restricted level of experimentation (k_α). This in turn increases the cost of intellectual property rights arising from (a) loss of second-stage innovations due to licensing failure ($x_\alpha \cdot q(k_\alpha)$); (b) lower average value of second-stage innovations ($v_2(k_2) - v_2(k_\alpha)$); while having no effect on the surplus from extra first-stage innovations under IP. As a result the welfare under weak IP rises relative to the welfare under IP and the marginal p must rise.

Similarly if the relative size of high value second-stage innovation compared to a low value one rises this (a) increases the 'optimal' level of experimentation (k_2) relative to the restricted level of experimentation (k_α) (b) directly increases the benefit of experimentation. This again increases the reduced experimentation cost and the licensing failure cost but reduces the surplus from second-stage innovations under IP. As a result welfare under weak IP rises relative to that under IP and the marginal p must rise.

This result then establishes the validity of the rest of Figures 2 and 3 and implies the following corollaries regarding how the optimal policy regime in relation to intellectual property rights varies in response to changes in the exogenous environment:

Corollary 4.8. *Reducing experimentation costs make it more likely that a freer (weak intellectual property rights) regime will be optimal.*

Proof. Follows from previous propositions as summarised in Figure 2.   □

Corollary 4.9. *Increasing the differential between high and low value second-stage innovations (which could be interpreted as experimentation becoming more important for product quality) makes it more likely that a freer (weak intellectual property rights) regime will be optimal.*

Proof. Follows from previous propositions as summarised in Figure 3.   □

5. Conclusion

In this paper we have shown how asymmetric information about the value of follow-on innovations, combined with intellectual property rights such as patents, can result in licensing failure and hold-up. Presenting the policy decision as a choice between having or not having intellectual property rights, we have shown that, in contrast to parts of the previous literature, in some circumstances it may be optimal not to have intellectual property rights. For whilst intellectual property rights help transfer income from second-stage to first-stage innovators they can also lead to licensing failure and hold-up with a resulting reduction in second-stage innovation.

The first part of the paper presented a basic model, which illustrated the already familiar trade-off when strengthening IP between reduced second-stage innovation as a result of licensing failure and increased first-stage innovation. This formed the basis for the model of cumulative innovation and experimentation which is the focus of this paper. Here, the existence of a perfect Bayesian equilibrium was shown along with the fact that, stronger IP, may have additional costs beyond those from licensing failure in the form of reducing experimentation below be socially optimal level.

Therefore, in addition to the basic trade-off mentioned above between more first-stage innovations and fewer second-stage ones, there is the additional factor: the average value of second-stage innovations have a lower average value due to a lower level of experimentation. Furthermore, this 'reduced experimentation cost' may be substantially larger than the costs of licensing failure. Examining the comparative statics, it was shown that, the lower

the cost of experimentation and the greater the differential between the low and high values of second-stage innovations, the more likely it is that a regime with weak intellectual property rights, in which second-stage firms did not have to license, was socially optimal.

Thus, technological change which reduces the cost of encountering and trialling new 'ideas' should imply a reduction in the socially optimal level of intellectual property rights such as patents and copyright. A perfect case of such technological change in recent years can be found in the rapid advances in computers and communications. These advances have, for example, dramatically reduced the cost of accessing and re-using cultural material, such as music and film, as well as greatly increasing the number of 'ideas' that a software developer can encounter and trial. Concrete policy actions that could be taken in line with these conclusions include extending 'fair-use' (fair-dealing) provisions in copyright law to increase the degree of reuse that would be permitted without the need to seek permission and excluding software and business methods from patentability.

Finally, we should emphasize that there remains plentiful scope to improve and extend the present paper. For instance, it was assumed that the non-royalty income for the first-stage and second-stage innovator was unaffected by the intellectual property rights regime.[16] However this is unlikely to be the case and the model could be improved by the inclusion of the direct effect of no (or weaker) intellectual property rights on the revenue of the first-stage (and second-stage) innovator.

It would also be useful to extend the analysis to the case of a continuous distribution of innovation values, as well as to investigate the consequences of making experimentation costs a function of the intellectual property rights regime. It would also be valuable to examine what occurs when the structure of innovation is more complex, for example by having second-stage inventions incorporate many first-stage innovations (a componentized model) or having heterogeneity across innovations with some developments used more than others. Finally, one of the most important extensions would be to properly integrate transaction costs into the analysis. Transaction costs relating to both the acquisition of information and the execution of contracts are significant and without them we lack a key element for the furtherance of our understanding of the process of innovation both in this model and in general.

---

[16]As discussed in detail above, while we do allow for business stealing between the first and second-stage innovators we do not allow for general rent dissipation from wider product market competition.

A. Proof for Basic Model

A.1. Proof of Proposition 3.1.

Proof. We are considering only subgame perfect nash equilibria so we may begin at the final stage of the game and work backwards. Given a royalty level of r, at the final stage, a second-stage innovator of type X faces a payoff of $v_2^X - r$ if she invests and 0 if she does not. Thus, a second-stage innovator, seeking to maximize profits will invest if and only if $v_2^X \geq r$ (formally, they are indifferent if $r = v_2^X$. However if they do not invest when $v_2^X = 0$ there will be no equilibrium of the overall game).

Given this, by simple dominance and focusing on pure strategies, a first-stage innovator must EITHER (a) set a low royalty rate $r_L = v_2^L$ which will lead to investment by all second-stage innovations; OR (b) set a high royalty rate $r_H = v_2^H$ which will result in investment only by high value second-stage innovations. In the first case the payoff is r_L while in the second it is $(1 - q)r_H$. Thus, a low royalty rate should be chosen if and only if (assuming that if payoffs are equal a low royalty is chosen):

$$r_L \geq (1 - q)r_H \iff q \geq \frac{r_H - r_L}{r_H} = \alpha$$

Since any mixed royalty strategy must consist of some combination of r_L and r_H we have immediately that a proper mixed strategy is only possible when $r_L = (1 - q)r_H$, that is if $q = \alpha$.

Finally, total royalty income to a first-stage innovator is at least $r_L = v_2^L$. Thus, total net income for a low-value first-stage innovator is at least $v_1^L + r_L = v_1^L + v_2^L > 0$ (by assumption) – and net income for a high-value first-stage innovator is obviously greater. Hence both types of first-stage innovator will invest. □

B. Proofs for Experimentation Model

We begin with a set of preliminary propositions which detail the players best responses.

Proposition B.1 (Second-stage innovator's investment strategies). *A second-stage innovator with value v_X facing a royalty of r will invest if and only if $v_X \geq r$.*

Proof. Just as in the original model second-stage innovator's move with full knowledge of all variables. In this case an innovator of type X invests if and only if net profits from

investing, $v_X - r - k\tau$ are greater than $-k\tau$ the payoff from not investing (experimentation costs are sunk). Hence the investment strategies are the same as in the basic model: a second-stage innovator invests if and only if $v_X \geq r$.   □

Proposition B.2 (First-stage Best-Response Royalty). *Under the strong IP regime, a first-stage innovator, whose belief about the experimentation level is given by the cdf $F(k)$ and where $\bar{q} = \mathbb{E}_F(q(k))$, will set a royalty of the form:*

$$
r(k) = \begin{cases} r_L = v_2^L, & \bar{q} > \alpha \\ r_H = v_2^H, & \bar{q} < \alpha \\ \text{mixed strategy } (r_H, r_L) \text{ with prob } (x, 1-x), x \in [0, 1], & \bar{q} = \alpha \end{cases}
$$

Proof. Given a first-stage innovator believes $F(k)$, the expected probability that a second-stage firm is low value is $\mathbb{E}_F(q(k)) = \bar{q}$. By subgame perfection a first-stage innovator knows that, once a seond-stage firm discovers its type, its best response to a given royalty will be as stated in Proposition B.1. In particular, if the royalty rate is set to be less than or equal to the second-stage low value (v_2^L) all second-stage innovators will license, if a royalty is above this but less than or equal to the second-stage high value (v_2^H) then only high value firms will license $(1 - \bar{q}$ of them) and if the royalty is higher than this no second-stage firms will license. Then, letting $G(r)$ be the cumulative distribution function over royalties representing the first-stage innovator's mixed strategy, the expected payoff to a first-stage innovator is:

$$
\Pi_1(G(r)) = \int_0^{v_2^L} r \cdot dG(r) + (1 - \bar{q}) \int_{v_2^L}^{v_2^H} r \cdot dG(r) + 0 \cdot \int_{v_2^H}^{\infty} r \cdot dG(r))
$$

Maximizing with respect to $G(r)$ immediately gives that, just as for the basic model, an optimal mixed strategy can only consist of some combination of the pure strategy $r_L = v_2^L$ and the pure strategy $r_H = v_2^H$. Let us suppose that these two pure strategies, r_H, r_L, are played with probability $x, 1 - x$ respectively. Revenue from royalties is then:

$$
r_L(1 - x) + (1 - \bar{q})r_H x = r_L + x \cdot ((1 - \bar{q})r_H - r_L)
$$

Maximizing revenue requires $x = 0$ if the term in brackets is less than zero, $x = 1$ if the term in brackets is greater than 0, and allows any value of x if the term in brackets is zero. By the definition of α (see above) these conditions correspond precisely to $\bar{q}$ (the

	$r \leq r_L$	$r_L < r < r_H$	$r \geq r_H$
$\Pi(k)$	$-k\tau - r + q(k)v_2^L + (1 - q(k))v_2^H$	$-k\tau + (1 - q(k))(v_2^H - r)$	$-k\tau$

TABLE 4. Payoff for Second Stage Innovator

expected probability of a low value innovation) being less than, greater than or equal to α. Hence, the first-stage innovator's royalty response as a function of their belief about the level of experimentation is of the form stated. □

Definition of k_α: If a first-stage innovator believes second-stage innovators all play the same pure strategy, k, then we can replace the conditions of the form $\bar{q} <, =, > \alpha$ with the condition that $k >, =, < k_\alpha$ (note the inversion of ordering), where as stated above constant k_α, is the experimentation level such that $q(k_\alpha) = \alpha$.

Proposition B.3 (Second-Stage Experimentation Level). *Under a strong IP regime the second-stage innovators best response to a royalty of r, including 'composite' royalties of the form $r = xv_2^H + (1 - x)v_2^L, x \in [0, 1]$ (that is mixed royalty with r_H played with probability x), is as follows:*

$$k = \begin{cases} k_2, & r \leq r_L = v_2^L \\ k_r, & r_L < r < r_H \\ 0, & r \geq r_H = v_2^H \end{cases}$$

where k_r is defined implicitly by:[17]

$$q'(k_r) = \frac{-\tau}{v_2^H - r}$$

And k_2 is given as follows:[18]

$$k_2 = k_{r_L} = k_{v_2^L} \Rightarrow q'(k_2) = \frac{-\tau}{v_2^H - v_2^L}$$

Proof. Using the optimal investment stage determined in Proposition B.1, for a given experimentation level k, payoffs as a function of the royalty levels are as in Table 4.

Suppose second-stage innovator plays a strategy given by the cdf $F(k)$ and a first-stage innovator sets a royalty defined by a cumulative distribution function $G(r)$. Then the

---

[17]If $q'(0) > -\infty$ then for values of r sufficiently close to $r_H = v_2^H$ this equation will have no solution. In such cases define $k_r = 0$.

[18]We use the subscript 2 because this is the level of experimentation undertaken in the case where both types of second-stage innovators find it worthwhile to invest.

payoff to a second-stage innovator is as follows (where expectations are taken with respect to F and q is short for $q(k)$):

$$\Pi_2(F(k)) = \mathbb{E}\left(-\tau k + \int_0^{r_L} qv_2^L + (1-q)v_2^H - rdG(r) + \int_{r_L}^{r_H}(1-q)(v_2^H - r)dG(r) + \int_{r_H}^{\infty} 0\, dG(r)\right)$$

$$= \mathbb{E}\left(-\tau k - q\{G(r_H)v_2^H - G(r_L)v_2^L - \int_{r_L}^{r_H} rdG(r)\} + G(r_H)v_2^H - \int_0^{r_H} rdG(r)\right)$$

Claim: Second-stage innovators play pure strategies.

Proof: q is convex so $-q$ is concave. Suppose we have a mixed strategy $F(k)$ with $\mathbb{E}(k) = \bar{k}$ then $-\bar{q} = \mathbb{E}(-q(k)) \leq -q(\bar{k})$ with equality if and only if $F(k)$ is a point distribution (i.e. corresponds to a pure strategy). Substituting:

$$\Pi_2(F(k)) = \mathbb{E}_F(-\tau k + G(r_H)v_2^H - q(G(r_H)v_2^H - G(r_L)v_2^L) - \int_0^{r_H} rdG(r)$$

$$= -\tau \bar{k} - \bar{q} \cdot (+\text{ve}) + \text{const}$$

$$\leq -\tau \bar{k} - q(\bar{k}) \cdot (+\text{ve}) + \text{const}$$

(With equality iff and only if $F(k)$ is a point distribution with $k = \bar{k}$ with probability 1). Thus for any properly mixed strategy $F(k)$ we can always achieve a higher payoff by playing the pure strategy $\bar{k} = \mathbb{E}(k)$. $\qquad\square$

Thus, in what follows we may confine our attention to pure strategies k. Returning to the payoff function we first note that if royalty (or royalties in a mixed strategy) are all greater than r_H (formally the support of $G(r)$ lies entirely above r_H) then the optimal experimentation level is zero ($\Pi_2(k) = -k\tau$).

When this is not the case we have the first order condition is:[19]

$$q'(k) = \frac{-\tau}{G(r_H)v_2^H - G(r_L)v_2^L - \int_{r_L}^{r_H} rdG(r)}$$

For ease of reference define S as the denominator in the previous equation. We shall look at several special cases as follows:

---

[19]The second order condition, $\Pi'' \leq 0$, is easily checked: $\Pi'' = -q''(k) \cdot (+\text{ve}) < 0$ since, by assumption, $q''(k) > 0$.

(i) $r \leq r_L$. Then $G(r_H) = G(r_L) = 1$ and we have $S = v_2^H - v_2^L$. The profit-maximizing k therefore equals k_2 where (as defined above):

$$q'(k_2) = \frac{-\tau}{v_2^H - v_2^L}$$

The intuition here is simple: both firms always invest and pay the royalty. Thus, in terms of the payoff, experimentation will only affect the value type and the experimentation level will be chosen so that the marginal gain in terms of lower costs, $q'(k)(v_2^H - v_2^L)$, equals the marginal experimentation costs, τ.

(ii) $r_L < r < r_H$. Here $G(r_L) = 0, G(r_H) = 1$ and we have $S = v_2^H - r$ and the optimal $k \equiv k_r$ solves:

$$q'(k_r) = \frac{-\tau}{v_2^H - r}$$

(iii) r_H played with probability x and r_L with probability $(1 - x)$. Then $G(r_L) = (1 - x), G(r_H) = 1$. Define the 'composite' royalty $r = xr_H + (1-x)r_L = xv_2^H + (1-x)v_2^L$ then we have $S = v_2^H - (1 - x)v_2^L - xr_H = (1 - x)(v_2^H - v_2^L) = v_2^H - r$. So the optimal experimentation level is $k \equiv k_r$ where r is the composite royalty.

$\square$

B.1. Proof of Proposition 4.2.

Proof. We will solve for a subgame perfect Bayesian nash equilibrium by recursing backwards through the game.

In previous propositions we have already derived the best-response correspondences (where the royalty best-response is defined in terms of *beliefs* about experimentation rather than the actual experimentation level). We have also shown second-stage firms will always play a pure strategy (i.e. choose a single experimentation level). Furthermore, at the experimentation stage all second-stage firms are the same, hence all second-stage firms will choose the same pure experimentation strategy. Thus, a first-stage innovator's beliefs (to be consistent) must be single-valued and we may rewrite the royalty best-response

correspondence in terms of their belief as to the experimentation level (k):[20]

$$r(k) = \begin{cases} r_L = v_2^L, & k < k_\alpha \\ r_H = v_2^H, & k > k_\alpha \\ \text{mixed strategy } (r_H, r_L) \text{ with prob } (x, 1-x), x \in [0,1], & k = k_\alpha \end{cases}$$

Case 1: $k_2 \le k_\alpha$. There are three possibilities for the beliefs of a first stage innovator regarding the experimentation level of second-stage firms:

(i) $k > k_\alpha$. Hence the first-stage innovator would set a high royalty rate. Then second-stage innovator's best response is $k = 0$ and beliefs will be inconsistent. Thus, there cannot be an equilibrium with such beliefs.

(ii) $k < k_\alpha$. In this case the best response of a first-stage innovator is to set a low royalty (r_L) in which case second-stage firm must choose a experimentation level $k = k_2$. Thus, for beliefs to be consistent, a first-stage innovator must believe $k = k_2$ and the equilibrium is as claimed.

(iii) $k = k_\alpha$. In this case a first-stage innovator's best response correspondence consists of all mixed strategies: r_H with probability x, r_L with probability $1-x$ for $x \in [0,1]$. Now a second-stage innovator (if behaving optimally) never samples above the level k_2 and will sample strictly below k_2 if the first-stage innovator plays any strategy in which r_H is played with positive probability. Hence if beliefs are to be consistent we must have (a) $k_2 = k_\alpha$ and (b) $x = 0$ (i.e. a low royalty is always set). In such a case the equilibrium is again as claimed.

Case 2: $k_2 > k_\alpha$. There are three possibilities for the beliefs of a first stage innovator regarding the experimentation level of second-stage firms:

(i) $k > k_\alpha$. Just as in the first case this leads to inconsistent beliefs and so cannot be an equilibrium.

(ii) $k < k_\alpha$. In this case the best response of a first-stage innovator is to set a low royalty (r_L) in which case second-stage firm must choose a experimentation level $k = k_2$. But $k_2 > k_\alpha$. Thus, beliefs will be inconsistent and this cannot be an equilibrium.

---

[20]At the experimentation stage all second-stage firms are the same and their best-response correspondence is single-valued. Hence all second-stage firms must have the same experimentation strategy and a first-stage innovator's belief

(iii) $k = k_\alpha$. In this case a first-stage innovator best response correspondence consists of all mixed strategies: r_H with probability x, r_L with probability $1 - x$ for $x \in [0, 1]$. Denote the corresponding composite royalty by $r(x) = xr_H + (1-x)r_L$. Then for an equilibrium (with consistent beliefs) we must find an x such that the best-response experimentation level equals k_α. Formally, using the notation of Proposition B.3 we must find an x such $k_{r(x)} = k_\alpha$. The best response experimentation level is defined implicitly by:

$$q'(k) = \frac{-\tau}{(1 - x)(v_2^H - v_2^L)}$$

Since $q' < 0$ we have, denoting $k(x)$ as the implicit solution as a function of x, that $k'(x) < 0$ (intuitively a higher average royalty lowers experimentation). Since $k(0) = k_2 > k_\alpha$ and that $k(1) = 0$ (as $x \to 1$ the RHS of the above takes arbitrarily large negative values), by the intermediate value theorem and the monotonicity of $k(x)$, there must exist a unique $x_\alpha \in (0, 1)$ such that $k(x_\alpha) = k_\alpha$. Replacing $q'(k)$ by $q'(k_\alpha)$ and rearranging we have as claimed that:

$$x_\alpha = 1 - \frac{\tau}{-q'(k_\alpha)(v_2^H - v_2^L)}$$

First-stage innovators investment strategy: finally as with our basic model first-stage innovators of both types invest because with royalty income net profits will be non-negative. □

B.2. Proof of Proposition 4.3.

Proof. Analogously to the low royalty case in the basic model, in this situation all second-stage innovators invest so (a) there is no licensing failure (b) second-stage firms sample at the optimal level (k_2). At the same time, intellectual property allows some first-stage innovators to engage in production who wouldn't be able to do so otherwise. Hence a strong IP regime will deliver higher welfare.

Formally, the welfare difference between the IP and WIP regime is net surplus associated with the p extra first-stage innovations that occur under IP:

$$p((v_1^L + r_L) + (v_2(k_2) - r_L))$$

Both the first term (by the assumption that the royalty is sufficient to allow production) and the second (since second-stage innovators are making non-negative profits) are positive. Hence, if $p > 0$ the sum is positive and welfare is higher with intellectual property. $\square$

B.3. Proof of Proposition 4.4.

Proof. In this case comparing the IP to the no/weak IP regime we have the following differences:

(+) Under IP there are (p) extra first-stage (and dependent second-stage) innovation because the royalty income allows some first-stage innovators to produce who would not otherwise:

$$p \; \underbrace{(v_1^L + v_2(k_\alpha) - x_\alpha q(k_\alpha) v_2^L)}_{\text{surplus per extra first stage innovation}}$$

(-) For the $(1 - p)$ first-stage innovations that occur under both IP and no/weak IP there are fewer associated second-stage innovations due to licensing failure (licensing failure cost) and the innovations are of lower average value due to reduced experimentation (reduced experimentation cost):

$$-(1-p)(\; \underbrace{(v_2(k_2) - (v_2(k_\alpha))}_{\text{Reduced Experimentation Cost}} \; + \; \underbrace{x_\alpha q(k_\alpha) v_2^L}_{\text{Licensing Failure Cost}} \;)$$

A strong IP regime is optimal compared to a weak IP (WIP) if the first effect is larger than the second (and vice versa):

$$p(v_1^L + v_2(k_\alpha) - x_\alpha q(k_\alpha) v_2^L) - (1-p)(v_2(k_2) - v_2(k_\alpha) + x_\alpha q(k_\alpha) v_2^L) \geq 0$$

$$\Leftrightarrow p \geq p^m \equiv \frac{(v_2(k_2) - v_2(k_\alpha)) + x_\alpha q(k_\alpha) v_2^L}{(v_2(k_2) - v_2(k_\alpha)) + x_\alpha q(k_\alpha) v_2^L + ((v_2(k_\alpha) - x_\alpha q(k_\alpha) v_2^L) - (-v_1^L))}$$

Where p^m has been defined as the probability of a low value first-stage innovation which leaves one indifferent between having and not having intellectual property rights. $\square$

B.4. Proof of Proposition 4.7.

Proof. Define:

$$S = \text{Higher Experimentation Cost} = (v_2(k_2) - v_2(k_\alpha))$$

$$H = \text{Licensing Failure Cost} = x_\alpha q(k_\alpha) v_2^L$$

$$E = \text{Surplus per Extra Stage 1} = v_2(k_\alpha) - x_\alpha q(k_\alpha) v_2^L - (-v_1^L)$$

Then,

$$p^m = \frac{S + H}{S + H + E}$$

Examining the differentials of S, H, E we have:

$$\frac{dS}{d\tau} = \frac{\partial}{\partial \tau}(v_2(k_2) - v_2(k_\alpha)) + \frac{dv_2(k_2)}{dk_2}\frac{dk_2}{d\tau} - \frac{dv_2(k_\alpha))}{dk_\alpha}\frac{dk_\alpha}{d\tau}$$

$$= (-) + (+ \cdot -) + (+ \cdot 0) = -$$

$$\frac{dH}{d\tau} = \frac{dx_\alpha}{d\tau}(\cdots) + (\cdots)\frac{dk_\alpha}{d\tau} = (- \cdot +) + 0 = -$$

$$\frac{dE}{d\tau} = \frac{dE}{dk_\alpha}\frac{dk_\alpha}{d\tau} = (\cdots) \cdot 0 = 0$$

Similarly,

$$\frac{dS}{dv_2^H} = +$$

$$\frac{dH}{dv_2^H} = +$$

$$\frac{dE}{dv_2^H} = -$$

For the last equation note, that by definition of k_α, $v_2(k_\alpha) = (1 + q(k_\alpha))v_2^L - k_\alpha \tau$ and that for $k < k_2, v'(k) > 0$ so that:

$$\frac{dv_2(k_\alpha)}{dv_2^H} = \frac{\partial v_2(k_\alpha)}{\partial v_2^H} + v'(k_\alpha)\frac{dk_\alpha}{dv_2^H} = 0 + (+ \cdot -) = -$$

Putting these derivatives together with the derivative of p^m with respect to S, H, E we have the required result. □

REFERENCES

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.

Howard F Chang. Patent Scope, Antitrust Policy, and Cumulative Innovation. *The RAND Journal of Economics*, 26(1):34–57, 1995. ISSN 07416261.

Iain Cockburn. Blurred Boundaries: Tensions Between Open Scientific Resources and Commercial Exploitation of Knowledge in Biomedical Research, 2005. mimeo.

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

R. Eisenberg and M. Heller. Can Patents Deter Innovation? The Anticommons in Biomedical Research. *Science*, 280(5364):690–701, 5 1998.

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

Jerry Green and Suzanne Scotchmer. On the Division of Profit between Sequential Innovators. *Rand Journal of Economics*, 26(1):20–33, 1995.

Bronwyn Hall and Rosemarie Ziedonis. The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979-1995. *Rand Journal of Economics*, 32(1):101–128, 2001.

Carmen Matutues, Pierre Regibeau, and Katharine Rockett. Optimal Patent Design and the Diffusion of Innovations. *Rand Journal of Economics*, 27(1):60–83, 1996.

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

Ted O'Donoghue, Suzanne Scotchmer, and Jacques Thisse. Patent Breadth, Patent Life, and the Pace of Technological Improvement. *Journal of Economics and Management Strategy*, 7:1–32, 1998.

Ariel Pakes. Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica*, 54(4):755–784, July 1986. ISSN 00129682.

Arnold Polanski. Is the General Public Licence a Rational Choice? *The Journal of Industrial Economics*, 55(4):691–714, Dec 2007.

Mark Schankerman. How Valuable is Patent Protection? Estimates by Technology Field. *The RAND Journal of Economics*, 29(1):77–107, 1998. ISSN 07416261.

Mark Schankerman and Ariel Pakes. Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period. *The Economic Journal*, 96(384):1052–1076, December 1986. ISSN 00130133.

Suzanne Scotchmer. Protecting Early Innovators: Should Second-Generation Products be Patentable? *The RAND Journal of Economics*, 27(2):322–331, 1996. ISSN 07416261.