The Relationship Between Intellectual Property Law and Competition Law: An Economic Approach

Pierre Régibeau and Katharine Rockett
University of Essex and CEPR
Revised, June 2004

Abstract

This paper presents an economic analysis of the relationship between Intellectual Property (IP) Law and Competition Law. Contrary to some of the recent debate, our analysis emphasises the separation of IP Law and Competition Law: IP law should concern itself with assigning and defending intellectual property rights, while Competition Law should concern itself with the use of those rights. This separation extends to the enforcement of the law as well, where we argue that once property rights have been assigned, no further distinction based on intellectual or non-intellectual property should be made. While the IP/Competition Law interface has some specificity due to the types of behaviours that tend to arise more frequently where IP is concerned, we argue for a set of principles for Competition Policy that include restraint, a commitment not to revisit *ex post* the rights granted by IP law, and a commitment to make large changes in property right regimes only when very large changes in *ex post* regulation occur.

Introduction

The purpose of this paper is to present an economic analysis of Intellectual Property Right Law and its relationship with Competition Policy. The relevant economic literature on this subject is enormous and complex. Here, we will strive for simplicity, trying to extract the main concepts and proposing simple principles that might help to guide the application and design of both Intellectual Property and Antitrust laws. While our analysis does not account for every single aspect of Intellectual Property Law or every single competitive situation, we do believe that the analysis does derive useful general principles.

The overriding thesis of our paper will be the separation of Intellectual Property (IP) and Competition Law. This separation will apply to the design of the law: IP law should limit itself to properly *assigning* and *defending* property rights while Competition Law should be concerned with the *use* of such property rights. More precisely, Competition Law should be concerned only with the use and abuse of property rights that are sources of monopoly power. This principle of separation also applies to the enforcement of the law. The main theme here is the equality of treatment of various sources of monopoly power, i.e. of the use of various property rights. We will argue that once property rights of various types have been properly assigned, there is no reason for Competition Policy to further distinguish between the sources of monopoly power. In particular, there is no need to treat monopoly power based on IP as "special" because of some supposedly unique characteristics such as its important to innovation or the public good nature of information. This is taken into account already in the special types of property rights that apply to intellectual property.

We develop our main thesis in two parts. The first part of the paper outlines the Economics of property rights in general as well as the Economics of intellectual property. In this part of the paper, we argue that the Economics literature broadly supports a design of intellectual property protection that looks similar to the system already in place, taking into account the special features of intellectual property. In the second part of the paper, we use this to make our argument for independence. We

summarise our position in a set of principles for Competition Policy that include restraint, a commitment not to revisit *ex post* the rights granted by IP law, and a commitment not to make large changes in property right regimes only when very large changes in *ex post* regulation occur.

This main thesis does not mean that the IP/Competition Law interface does not have some specificity. We investigate this issue in the third part of the paper. We find that there are special concerns in the areas of mergers, licensing and cross-licensing, patent pools, grant-backs, various practices that extend the legal patent monopoly beyond the life of the patent, interfaces and interoperability, umbrella branding, and compulsory trademark licensing. Largely, these come from special behaviours that are rarely observed in practice when market power comes from the ownership of other types of exclusive assets.

Part I. Economic Analysis of (Intellectual) Property Rights

We first present an overview of the economic analysis of property rights. We ask only later if and how intellectual property rights might be distinct. We find that there are two main differences between intellectual property and real property: intellectual property has strong public good characteristics and tends to generate significant amounts of socially useful information, making diffusion of information an important concern. These differences help to understand the distinct property regimes that apply to different types of intellectual and real property. In particular, we concentrate on two aspects of property rights: the duration for which they are imparted and the "scope" of the rights. We follow the economics literature's analysis of the economically optimal duration and scope of rights, finding that the current system does not differ greatly from the system that would be recommended by the existing economics literature. Patents, copyrights and trademarks are considered in turn.

1. Property Rights and Efficiency

From an economic point of view, property rights are a necessary condition to achieve *efficiency*. The economic concept of efficiency relies on the notion of *opportunity cost*. The opportunity cost of using any resource is the return that this resource would have obtained in its best possible alternative use. Hence the opportunity cost of using a piece of land to build new flats might be the net benefits that would have accrued to the community if the land had been developed as a park. It is customary to distinguish between *static* and *dynamic* efficiency.

A. Efficiency

Static efficiency is attained when consumers and producers make their decisions, taking into account the true opportunity cost of the resources involved. It is convenient to distinguish between two aspects of static efficiency. Consider first the market for a single product in isolation. *Static* efficiency requires that the price of the good be equal to the marginal (opportunity) cost of producing it. This ensures that consumers take the appropriate cost of production into account when deciding how much of the product to purchase, so that they purchase neither too little nor too much of the good. If price were above marginal cost – as it is would be in the presence of market power – then consumers would purchase too little of the good. The additional satisfaction that consumers would have enjoyed from consuming the units that they would have bought had the price been set at marginal cot is the static social cost of monopoly. It is called the deadweight loss (see figure 1).[1] Once many sectors are considered simultaneously, the fact that the relevant cost of production is the marginal *opportunity* cost takes its true meaning. As the cost of using a resource is equal to the benefits that it would have generated if it had been employed in another sector, the allocation of resources should be such that their marginal returns are equated across

[1] We have assumed that there are no direct *externalities*. Direct externalities are cost (or benefits) imposed by one economic agent on others. Examples include pollution – *you* breath the exhaust of *my* car – or the fact that you get some enjoyment of living in a well-maintained neighbourhood – *you* enjoy some of the benefits from *my* maintenance work. Direct externalities drive a wedge between private opportunity cost – i.e. the cost taken into account by the agents making the decisions – and social opportunity cost – the cost that should be taken into account from the point of view of society as a whole. With externalities, the principles discussed above still hold as long as opportunity cost is understood as the 'social opportunity cost' that incorporates all relevant external effects.

sectors. This aspect of static efficiency is sometimes referred to as *allocative* efficiency.

There is no universally accepted definition of *dynamic* efficiency. For our purpose, it suffices to say that it relates to any kind of investment decision. This concept of 'investment' must be understood broadly. While it clearly covers the development of improved machines, products or methods of production (i.e. "Research & development), it also includes the creation of physical assets (e.g. new plants, or the clearing of new arable land) as well as their maintenance.

As for static efficiency, let us first consider one investment project in isolation. Dynamic efficiency requires that the project be undertaken if – and only if – its social benefits exceed the opportunity cost of the resources invested. In other words, inefficient investment decisions can occur when the stream of private revenues generated by the investment is not equal to its social benefits. Even in the absence of direct externalities, there is no reason to believe that private investment choices will generally be efficient, as two opposing forces are at play. Consider a firm introducing a new product. On the one hand, it will not usually be able to capture all of the benefits that this new product creates for consumers. This is clearest when the firm charges the same price for each unit of the products as every consumer purchasing the good enjoys benefits that are *at least* as high as the price demanded. If consumers differ in their enjoyment of the good but pay the same price, some of them must inevitably obtain benefits that are strictly higher than the price paid. Hence this *appropriability* effect leads to insufficient investment as the firms cannot capture all of the value that they create. On the other hand, the new product is likely to reduce the sales of older existing products and the profits that other firms obtain from these sales. In other words, the introduction of the new product imposes a loss on other firms that is not properly considered by the innovating firm. This *business stealing* effect leads to excessive investment.

Once many sectors are considered simultaneously, dynamic efficiency also requires that the allocation of resources between different investment projects be correct. This requires that the (marginal) social rate of return of investment be equalised across sectors. There is unfortunately little hope that such efficiency will

always obtain, for two main reasons. Firstly, there are no reasons to believe that the appropriability effect and the business stealing effect discussed above will have the same relative strength across all sectors of activity. This means that, even in the absence of any policy intervention, one would expect over-investment in some sectors and under-investment in others. Secondly, differences in private returns to investment can result from public policy measures. In particular, an uneven application of competition law across sectors or across various sources of monopoly power (e.g. intellectual property versus other forms of property rights) would create artificial differences between private returns to various types of investments.

B. Property Rights

From an economic point of view, property rights are assigned for four main reasons.

1. To maintain peace and order. Together with other aspects of the legal system, legally enforceable property rights help minimise physical violence – and the associated destruction of economic resources - aimed at securing the control of assets that are sources of economic rewards. This factor clearly applies to all types of property, 'real' or 'intellectual'.

2. To assign decision rights. From our discussion of static efficiency, it should be clear that an efficient allocation of resources cannot obtain unless someone (individual or collectivity) has the right to decide how economic assets are used. However, the consequences of this principle are rather different depending on whether it is applied to 'real' or to 'intellectual' property. This is because most forms or real property are seen as *private goods* while intellectual property is generally thought to a *public good*. Private goods are characterised by *rivalry* in usage, i.e. they cannot be used by more than one economic agent at the same time: we cannot both eat the same apple. Because of this, it is generally optimal to let a single agent decide how the good ought to be used. Public goods, on the other hand, are such that usage by one agent does not preclude usage by another: unless the beach is very busy, my view of the sea does not 'crowd out' your view of the sea. As long as there is no rivalry at all in the consumption of a good, it is optimal to let as many agents as possible have access to

it. This is the case with knowledge: my using Pythagoras' theorem does not in any respect restrict your own access to it. Hence, once knowledge has been created, the best possible policy is one of free access, making exclusive property rights undesirable. To summarise then, the need to assign decision rights is one of the reasons for allocation property rights on *private* goods but not on *public* goods such as most forms of intellectual property.

3. To reward investment. This factor relates to dynamic efficiency. The idea is simply that no rational economic agent will incur the cost of investing in developing or maintaining property unless she his able to collect some corresponding reward. Hence, if sufficient investment is to be induced, investors must be given property rights over the fruits of their investment so that they can capture a significant proportion of the value that they create. Although this factor applies to all types of property, it is of special importance for assets whose development and/or maintenance require significant effort.

4. To favour the diffusion of information. Agents investing in assets might try to exploit them 'secretly', expending effort to prevent others from gaining information about the asset. Hence a manufacturing firm might change its production process unbeknownst to the outside world or a prospector might go to great length to exploit a gold mine without revealing its location. Such secrecy is socially undesirable for two reasons. Firstly, the resources spent on preventing information from leaking out to other agents are diverted from productive uses. Secondly, if information about the asset is useful in the ulterior development of other assets, secrecy reduces the pace of economic development by limiting the information available to other investors.

This second aspect is an issue of dynamic efficiency. Although similar in spirit, it differs from the third factor discussed above, whereby property rights were required to provide a sufficient reward to investors. To see this, consider the case where an invention cannot be reversed-engineered. Property rights are not necessary in order to reward the inventor's investment since the invention can be commercialised successfully without any risk of imitation. On the other hand, the inventor will only reveal information about the innovation if sufficient protection is offered *and* if this protection is made *conditional* on the release of information. Without protection the

inventor would not only fear direct imitation of her innovation but also the faster emergence of new innovations making her own inventive step obsolete. Without conditionality, the inventor would have no incentive to reveal any information, regardless of the level of protection provided.

This function of property rights is clearly most important for assets that generate large amounts of information that is potentially useful for further investments. As should be obvious from the example above, investments in intellectual assets tend to produce such information abundantly. While one can think of similar examples for some real assets (e.g. prospecting for gold), this 'information' diffusion factor seems to be more relevant for intellectual property.

Considering these four factors together, we can identify two main differences between intellectual property and 'real' property: intellectual property has strong public good characteristics and tends to generate significant amounts of socially useful information, making diffusion of information an important concern. These differences help understand the distinct property regimes that apply to different types of intellectual and real property. For simplicity, we will concentrate on two aspects of property rights, the 'duration' for which they are imparted and 'scope' of the rights., loosely defined as 'what is actually protected by the rights'[2].

The scope of property rights on real assets is easily defined. There is usually little room for ambiguity: having rights on my house does not, for example, imply that I can prevent anyone from owning or using a copy of my house on the plot of land next to mine. Rights on real assets are also usually granted without an explicit time limit[3]. Going back to our economic analysis of property rights we can see that this design makes sense: because real property tends to be *rival*, it is efficient to have exclusive property rights assigned *throughout the useful life of the asset* (see factor 2 above). Finally, because diffusion of information is a lesser concern, rights on real

[2] We do not use the term 'scope' to define how the property right owner can – or cannot – *use* her asset. We focus instead on the exclusionary aspect of property rights. In this sense, the scope of the property right is the precise asset that others can be precluded from using.
[3] We refer here to the rights of the owner of the asset, not to the temporary rights that this owner might grant to others. These later rights are best seen as a use of the owner 's original property right.

assets can usually be obtained without having to explicitly divulge information about the asset and/or how it was developed.[4]

As we will discuss at some length in the following sections, the scope of intellectual property is itself a rather complex concept. One should therefore not be surprised that a large chunk of intellectual property law is devoted to its definition and implementation. We also know that, with (some types of) intellectual property, diffusion of information is an important concern. The fact that property rights are often limited duration should therefore not be surprising.

2. Application to IPRs

In this section we apply our economic analysis of property rights to make two main points. Firstly we show that the factors identified above can help us to organise an apparently disparate economic literature on IP protection. This is quite useful, as this literature is essential to any economic analysis of the interface between IP law and competition law. Secondly, we argue that there is not much distance between the type of IPR protection that economic analysis would recommend and the actual structure of (most) IPR Laws. In particular economic analysis supports the existence of distinct protection regimes for patentable material, copyrightable material and trademarks. In fact, it is when IPR Law departs from the traditional assignment of property between these regimes – as in the case of software protection - that it sometimes runs counter to economic analysis.

As we have seen above, the two main factors explaining the need to grant property rights on intellectual assets are the *reward* effect, which induces investment by ensuring that the inventor can reap significant benefits from her innovation and the *information diffusion* or *disclosure* effect, whereby the inventor is 'bribed' into divulging useful information about her innovation. Both effects involve trade-offs between dynamic and static efficiency. The right to exclude others from using an

[4] There are of course exceptions such as the rights given to prospectors – from gold diggers to oil companies -, which are typically for a limited duration. One of the reasons for this is that, in most cases, the useful life of the asset is itself rather short. Such 'real' assets have other 'intellectual asset' characteristics such as the fact the scope of the rights is not always unambiguous (e.g. adjacent natural

intellectual asset for a given length of time only provides the rights-holder with an economic reward or 'bribe' if it provides some *monopoly power*. The more significant the monopoly power, the larger the reward but, also, the larger the loss in static efficiency due to the fact that prices rise above opportunity costs. The general principle guiding the design of intellectual property rights should therefore be to find the combinations of duration and scope that ensure a given level of reward to the inventor at the smallest possible cost in terms of induced inefficiencies.

a. Patents

The economics literature on the design of the patent system is very diffuse. Different conceptions of how innovation occurs and different aspects of the patent system have led to a wide variety of analyses. Still, these disparate pieces can be fit broadly into the two effects described above. This section will briefly review some of the more important papers in the literature, attempting to fit the approaches of various authors into our simple framework. While this review is not exhaustive, it should give a feel for how the principles derived from the economic theory of property rights can be apply to design an efficient system of patent protection.

A first strand of the economic literature on patents assumes that there is a single potential innovation for which inventors compete. This set up rules out any role for our disclosure effect since further inventions are not considered. When we only consider the *length* of the period of protection, the reward effect also takes a very simple form: the longer the length of protection, the greater both the reward and the social cost from the resulting monopoly power. The aim of the earliest contribution to this literature was precisely to generate an optimal patent length, balancing the rewards as inducements to innovation against the deadweight loss generated by a grant of patent monopoly (see Nordhaus, 1969, and Scherer, 1972.)

This approach was extended to include the determination of both the length and the *breadth* of patent protection. The issue then is how to *structure* the reward to innovation, for a given reward size. As the analysis only relates to a single innovation,

---

gas fields might be linked geologically) or the fact that the very location of the find provides

the appropriate notion of breadth is quite straightforward: the breath of patent protection is increased by any measure that increases the cost of imitating the protected invention. The first contribution to this line of research was Tandon's analysis of compulsory licensing (Tandon, 1982). In his model, patent holders are obliged to license their intellectual property to all comers at a regulated royalty rate. The lower the royalty, the lower the effective monopoly power of the patent-holder and, therefore, the longer the period of protection required to ensure that inventors obtain an adequate reward on their investment. In other words, the regulated royalty rate plays the role of patent breadth. In Tandon's model, a very long patent (infinitely-lived, in fact), accompanied by a very low regulated royalty rate on compulsory licenses is optimal as this minimises the monopoly distortion per period while maintaining innovation incentives. In fact, Tandon points out that this result favouring length over scope of protection follows from the fact that social welfare is far more sensitive to the royalty rate in his model than to the length of time during which the royalty is paid.

The trade-off involved in Tandon's work is analysed in quite general terms by Gilbert and Shapiro (1990) who simply assume that, during the period of protection, greater breadth increases the firm's profit but decreases welfare. They then show that the crucial element is whether greater breadth increases profits *more rapidly* than the associated welfare deadweight loss. If it does, then a regime with broad protection for a short period of time is optimal. If it does not, then small breadth and long length are called for. While quite general, this principal is not especially useful if one does not know which way the comparison goes. For a relatively general case, where profits and welfare are both concave in output[5], Gilbert and Shapiro show that the most efficient way to reward innovation is to grant very long patents that are only just broad enough to attain the desired level of reward. They point out, however, that their results favouring long patent lives crucially depend on the assumption that patent breadth only affects price. Hence, the extent of substitution away from the patented product, and the associated deadweight loss, always increase with breadth[6].

---

information about the type of terrains where similar assets might be found.

[5] This is satisfied by relatively standard assumptions that demand and marginal revenue slope down while marginal cost slopes up and goods are homogeneous.

[6]Their results also require that there be no uncertainty about profits, that the underlying environment is stationary, and that no further improvements to the technology will occur. In the conclusion to their

Klemperer (1990) considers a somewhat different situation, where products are not homogeneous. In his paper imitators produce 'knock offs' that are of lower quality than the patented product. In this framework, broad patents create a large distortion from monopoly pricing of the patented good but narrow patents result in a suboptimal allocation where most consumers purchase the less desirable 'knock offs'. Klemperer shows that, depending on whether the effect of switching to knock off brands or switching out of the product class entirely dominates, the socially optimal manner of providing the inventor with a given reward can involve either narrow breadth with long length or large breadth with small length. Klemperer 's contribution must therefore be seen as showing that Gilbert and Shapiro 's conclusion that narrow patents are optimal is not robust . It only holds when consumers are similar along a particular dimension: they must all face similar costs of substituting to less-preferred varieties although they may differ in their costs of substituting out of the patent class. If consumers are similar to each other along other dimensions, this result may be completely overturned. Given that accurate information about the precise form of consumer preferences in each market is unlikely to be available, this portion of the literature does not yield strong policy recommendations toward either extreme in patent protection. The best route might therefore lie somewhere in between, arguing for a finite duration and moderate breadth[7].

We have assumed so far that, in principle, patents can be granted that confer some degree of monopoly power forever. The starting point of a second, more recent strand of the economics literature is that, in practice, the *effective lifetime of the patent* may be curtailed by subsequent innovation that supersedes the patented technology. Most of this literature still focuses on what we called the reward effect, but the trade-offs involved are more complex. This reflects the greater complexity of the notion of patent breadth itself. Economists distinguish between *lagging breadth* and *leading breadth*. [8] Lagging breadth is the protection granted against *imitation* while *leading*

---

paper, they note that their extreme conclusions on breadth could easily be overturned by relaxing either of these assumptions.

[7] Scotchmer and Maurer (2002) point out that a link can be made between the Gilbert and Shapiro definition of breadth and the definition in a related paper by Gallini (1992) that reaches the conclusion that some breadth is important to avoid imitation around the patent. They point out that Gilbert and Shapiro reach their conclusion of narrow breadth because they are concerned only with minimising the deadweight loss of the patent, whereas Gallini aims to minimise both deadweight loss and the wasteful cost of imitation.

[8] See O'Donohue (1998) and O'Donoghue, Scotchmer and Thisse(1998).

breadth refers to protection with respect to further *improvements*. The distinction between imitation and improvements can be made quite precise: Assuming that the initial innovation is sold at marginal cost [9], an imitation does not increase social welfare, while an improvement does. In other words, improvements 'add value' but imitations do not. Notice however, that, to qualify as an improvement a subsequent innovation does not need to be strictly better than the initial invention. It suffices that it be better *for at least some users*. In economic terms this means that improvements include both 'vertical' and 'horizontal' differentiation**.** While lagging and leading breadth are economic concepts, they do correspond to distinct aspect of legal doctrine. In legal terms, the strength of lagging breadth is defined by the doctrines of disclosure and enablement , while leading breadth is determined by the interpretations of "use of a technology", the doctrine of equivalents, and the doctrine of reverse equivalents.

An innovation starts off a research route, after which other innovations build on the first. In this sense, the entire benefit to society that follows from this stream of innovations can be attributed to the first inventor. It follows that a social planner, taking into account this entire stream of benefits, would have full incentives to undertake efforts to initiate entire research paths. Private firms may not, however, as their reward to innovation is limited by time, the "leading breadth" of the patent (in other words the quality improvements that would actually fall under the original patent), and the transfer of demand to improved products. In order to restore their incentives, then, this literature argues for broad patent protection.

This argument has taken various forms. In Green and Scotchmer (1995), they interpret the novelty requirement as setting the minimum quality improvement that will infringe a patent. O'Donoghue (1998) makes this more precise by saying that a patentability requirement, specifying the minimum innovative step that would receive patent protection, is a combination of novelty, nonobviousness, and utility. If an initial innovation is to be improved by another firm, the initial innovator and the improver can increase their joint profits by signing licensing agreements to share industry profits. For example, an agreement can be signed that assures the initial innovator will have sufficient reward to induce him to undertake the initial

[9] This benchmark is easily defined if the innovation consists of a new product. For process innovation, the equivalent assumption is that it is licensed to all.

innovation, and also assures that the improver will have sufficient reward to induce investment in the improvement. Varying the patentability requirement changes the relative bargaining positions of the initial and follow-on innovators in such licensing agreements by changing the rewards to firms if they fail to strike an agreement. As Green and Scotchmer's concern is to transfer more profit to the initial innovator, they argue for a stringent patentability requirement.

In fact, taken literally, Green and Scotchmer's argument implies that all future innovations beyond some "first" in a research stream should be judged unpatentable and infringing.[10] However, Denicolo (2000) points out that the sequential innovation line of reasoning assumes that no patent race occurs to generate innovation. Adding such a race to the model has two effects. First, there may be socially excessive amounts of innovation: because of the *business stealing effect* discussed above, firms racing for a patent may invest too much compared to what would be socially desirable. Once the possibility of socially excessive innovation is introduced, welfare is not necessarily improved by raising the reward of the first innovator as much as possible. Furthermore, it is possible to show that, when competition for the innovation is explicitly considered, leading breadth might best be relatively narrow for the initial innovation. The reason for this is that too large a breadth on the initial innovation limits the firms ' incentives to racing for improvement: if 'winning the race' means paying large royalties to the initial inventor, why bother? In other words, if the improvements are, effectively, reserved for the initial inventor then the lack of competition may result in under-investment in later stages of the inventive process.

It is also useful to understand that most of the papers in this literature assume a specific antitrust regime in the sense that particular licensing arrangements are assumed to be available to firms, or particular pricing policies are allowed. The optimal patent policy tends to change depending on the antitrust assumptions made. In Gilbert and Shapiro, for example, the breadth of the patent has the effect of raising price, but so would loose competition policy. Hence a more permissive application of competition policy would make it desirable to narrow the breadth of patent protection. In Green and Scotchmer (1995) licensing agreements that divide industry profits earned from the stream of innovation are allowed before the improving innovation

occurs. In fact, a liberal licensing policy is optimal as it increases the maximum reward that can be transferred to the first innovator. O'Donoghue, Scotchmer and Thisse (1998) also suggest that allowing collusive agreements for stimulating the flow of R&D might be beneficial. However, these conclusions are not robust for the reason discussed in the previous paragraph: once competition for the innovations is considered, increasing the reward of the first innovator as much as possible is no longer optimal.

O'Donoghue (1998) modifies the approach taken by Scotchmer and co-authors and assumes that the size of the innovative step is a choice variable for the innovating firm. In other words, O'Donoghue focuses not on how patentability affects the *bargaining* power of firms when innovations are improved in given steps, but how a stringent patentability requirement can induce firms to invest in bigger steps and so change the actual *type* of innovation that occurs. O'Donoghue argues that forcing firms to only patent bigger steps makes innovation more infrequent. As this increases the average incumbency period of an innovator (the *effective length* of protection), it increases the reward to innovation. This, in turn, stimulates further investment in research.[11] On the other hand, bigger inventive steps also mean that the latest innovation faces less of a competitive constraint from the previous generations of inventions, increasing the static welfare loss. The optimal patent regimes relies on significant – but finite – leading breadth to balance these two effects.

To summarise this second strand of the economics literature, there appears to be little robust argument for either of the extreme policies of very long, narrow patents or very short, wide patents. Some middle ground appears to be the more reasonable path that balances the need to compensate early innovators for the externality they generate in terms of stimulating future innovation, while both providing sufficient incentives for researchers to take those follow-on steps in a timely manner. Any more precise recommendations tend not to be general and so are not appropriate to a policy that is not tailored to particular industries.

---

[10] See Denicolo (2000)

[11] Eswaran and Gallini (1996) also investigate the effects of patents on the *type* of research conducted, focussing on whether product or process innovation will occur. In the absence of patents, too much product innovation occurs, as this tends to attenuate competitive pressures, while too little process innovation occurs as this tends to increase competitive pressures. This bias can be rectified by modifying the relative patent breadth appropriately on process and product innovations.

The streams of literature mentioned above do not generally take into account what we have called the information diffusion effect. Exceptions to this are Scotchmer and Green (1990) and Matutes, Regibeau and Rockett (1996). Both of these papers assume some kind of sequential innovation, but the emphasis is different. Scotchmer and Green identify the following trade-off. While a strong novelty requirement confers a large reward to innovation, because it reduces the substitutability between the patentable innovation and other technologies, it also means that relatively minor advances are not patentable. Since the informational spillovers from these minor patents would speed up the development of the field there a strong novelty requirement also has a social cost.

Matutes, Regibeau and Rockett (1996) focus on the patent protection that a basic innovation should be granted, when that innovation is likely to generate a stream of applications. In their model, the basic innovator may be tempted to "opt out" of the patent system, at least temporarily, in order to stockpile applications before applying for a patent[12]. This delay in applying for the patent reduces disclosure and so reduces the ability of other firms to benefit by developing applications themselves. The conclusion of the paper is that, by designing patents to grant a "limited license to hunt" for applications this wasteful delay can be reduced. Hence, patent protection should be extended to applications beyond the existing demonstrated usefulness of the product or process as specified in the claims in order to induce early disclosure of fundamental innovations, while preserving firms' incentive to innovate. The optimal scope, then, implies that inventors of basic innovations obtain protection on applications that have not yet been fully worked out, requiring a more lenient review of claims than is current practice.

[12] Referring to the United States, the *Gottschalk v. Benson* decision (S. Ct. 175 USPQ 673) established in 1972 that a pure algorithm is not patentable in the sense that one cannot pre-empt all uses of a mathematical formula. The *Diamond v. Diehr* case (In re Diehr, 1979-CCPA: 203 USPQ 44/Diamond v. Diehr 1981- S. Ct.: 209 USPQ 1) further clarified this by finding that the algorithm in question was patentable because it was applied to a particular process that was specified as part of the patent claims. More recent patents (such as 4,405,829) seem to have gone against this "narrow" interpretation. In terms of this model, a broad scope would be equivalent to allowing a patent on the formula itself in the *Benson* case, or on very loosely defined applications of the algorithm. A narrow patent would amount to following the philosophy of the *Diehr* decision that very few applications that are already well-developed would be protected by the patent for the period of its duration.

A smaller stream of literature considers whether the patent system might better be redesigned entirely. The more sophisticated and recent of these papers use the tools of mechanism design to completely re-work the system of rewarding innovation. Earlier papers adopt a simpler approach. Wright (1983), for example, considers whether a system of prizes in exchange for commissioned work would be more socially optimal. For example, such a system would not necessarily entail the deadweight loss that the patent system's conferral of monopoly generates. Unfortunately, Wright's system places strong informational requirements on governmental authorities so that they can correctly "pick winners". If one assumes that firms -- or individual inventors -- have better information than the government on the relative costs and benefits of innovation then some form of delegation, like patenting, might work better. In the more sophisticated strain, Kremer (1998) proposes an auction system to *supplement,* rather than replace, the existing patent system as a mechanism to stimulate innovation while reducing deadweight loss. He claims that this will solve some of the informational problems of Wright's framework by using the auction mechanism to allow private industry to "reveal" the true value of the innovation. He acknowledges, however, that collusion amongst the private firms in the auction could result in large compensation being paid to industry for innovations that had very little social value.

In sum, then, the reward function of patents has been investigated in quite a lot more detail than the disclosure function. The literature on optimal patent design has complemented this by investigating not only how protection should be structured to guarantee a given level of reward (whatever that might be), but also whether the patent system is optimal in a broader sense as a way of creating incentives to innovate. The "bottom line" on this stream of research, however, is that something that looks like the current patent system can be defended as optimal from an economics viewpoint. Within the current patent system, there is a relatively strong argument for broad lagging breadth, as well as for a finite *effective* patent length guaranteed through a combination of leading breadth and duration of patent protection.

3. Copyrights

Copyright laws are not meant to apply to the same type of material as patent laws. Very roughly, copyright laws seek to encourage and protect *creative expression*, while patent laws deal with *innovation*. While sometimes subtle, the difference between the two types of creative activities involved is not purely semantic.

Let us first consider our disclosure/diffusion effect. The 'ideas' or 'knowledge' contained in traditional copyrightable material are automatically revealed through the publication and (possible) sale of the product.[13] While the author can of course keep her work secret, she cannot commercially exploit it without revealing all of its socially useful content. The need to offer protection in order to induce the revelation of all relevant information is therefore smaller than in the case of patentable material. On economics ground, then, we would expect copyright protection to be 'weaker' than patent protection.

It is sometimes argued that the 'reward' effect is also of little relevance to copyrightable works as they are the result of a 'need to create' that is little influenced by economic incentives. If that were true then, the *economic* argument for offering any protection at all would be very weak. In fact, the (scarce) empirical evidence available suggests that the production of some types of copyrighted material does respond to economics incentives[14]. Hence, some amount of protection grounded in our 'reward' effect seems justified. But what form should this protection take?

Two characteristics of traditional copyrightable material are relevant. Firstly, the 'sequential' aspect of innovation, which was so important in determining the appropriate patent protection regime seems less acute. While creative works do sometimes trigger further waves of creation[15], this process seems somewhat more diffuse and harder to define than in the case of patentable innovations. This would make the determination of any significant 'leading breadth' somewhat hazardous. Furthermore, creative work is rarely made obsolete by its progeny: while a 'better mousetrap' makes the original invention useless, another impressionist painting

[13] We focus on the protection granted to *published* material. For a discussion of the protection of *unpublished* works, see Landes (1992).
[14] See Hui and Png (2002).
[15] One needs only think of the various 'schools' or 'movements' that have characterised literature or painting over the last two centuries.

hardly detracts from the value of earlier works from the same school.[16] One would therefore expect copyright law to offer little in terms of 'leading breadth', concentrating instead on *lagging breadth* – i.e. protection against *copying* – and length.

A second characteristic of importance is that, given the nature of the material it deals with, copyright law is strictly constrained by the desire not to infringe on freedom of expression. In terms of our analysis of property rights, this means that even protection against copying would be of limited scope as greater breadth might have an unacceptable stifling effect on public discourse. In other words, one would expect that protection against copying might not be absolute[17] and that the definition of what constitutes 'copying' would be rather narrow. As there are good reasons to keep both 'leading' and ' lagging' breadth narrow, copyright protection would naturally tend to rely more heavily on the *length* of the period of protection.[18]Overall then, based on the basic economic principles derived in section 1, copyright protection should be quite narrow but long. In that sense, actual copyright protection regimes closely resemble the optimal regime predicted by economic theory.[19]

On the other hand, our analysis also suggests that extending copyright protection to less traditional works might not be appropriate. In particular, using copyright law to protect creations that include *inventive steps* would not make sense. In the presence of such steps, issues of sequential innovation resurface, calling for a different protection regime. On economics ground therefore, there is a strong argument for protecting innovative software under patent law rather than under copyright law.

The recent economics literature on copyrights points out two further factors that differentiate copyrights from patents. Both factors refer to the traditional static trade-

[16] In fact, further works might even enhance the value of earlier works to the extent that they help 'cement' the legitimacy of the movement. In economic terms, such an effect would be seen as a form of 'network externality'.
[17] The doctrine of 'fair use' can be seen as an example.
[18] Notice that ,as obsolescence is not a significant issue, 'length' and 'effective length' essentially coincide.
[19] Hui and Png (2002) suggest that, in the US at least, copyrighted materials actually receive the maximum effective length of protection. They find that the extension of the length of protection following the Sonny Bono Copyright Extension Act of 1998, adding 20 years to the previous length, had no significant effect on the production of copyrighted movie material.

off between incentives to 'invent' and the welfare loss imputable to monopoly power. On the 'reward' side, several authors have shown that *copying does not necessarily hurt* the producer of copyrighted material. Although the precise mechanism involved varies, the key factor is that copying helps the monopolist *commit* to pricing strategies that she would like to implement. An example will suffice to illustrate this basic principle[20]. The first example relies on price discrimination. Suppose that the copyrighted good is characterised by *network externalities*, i.e. the valuation of the good for any given consumer increases with the number of other consumers who also have access to the product (or its copies).[21] To entice consumers to buy, the copyright holder would like to 'promise' to sell enough to create a large network of users. Unfortunately, this promise is not credible as potential buyers know that, once the firm has acquired a basic 'installed base' of consumers, it will find it optimal to increase prices and slow down the growth of the network. In this perspective, weaker copyright protection might help as it guarantees that a large number of copies will in fact be made available: this guarantee increases the willingness to pay of consumers making them more willing to purchase the product early on and at a higher price. The basic lesson from this strand of the literature is therefore that the reward obtained by the copyright-holder does not necessarily increase with the strength of copyright protection. In that sense, it is a further argument for granting weaker protection to copyrighted work than to patentable innovations.

On the other side of the traditional reward/monopoly trade-off, Novos and Waldman (1984) have shown that, under some conditions, *an increase in copyright protection can in fact increase static welfare*. In other words, greater protection might increase the copyright-holder 's reward (subject to the caveat above) without creating a corresponding deadweight loss. In their set- up, the marginal cost of physical production is the same for the original good and each of its illegal copies. However, copying incurs an additional cost that increases with the intensity of copyright protection. Illegal copies are supplied competitively at this augmented marginal cost. The respective market shares of the original product and its copies are such that the marginal customer is indifferent between the two options. This means that this consumer is willing to incur a copying cost equal to the price charged for the

[20] Other examples rely on a price-discrimination mechanism. See Takeyama (1997) or Bakos and al. (1999).

original product. Because the copyright-holder charges a price in excess of its own marginal cost, this means that the marginal cost incurred by the marginal 'copying' customer is higher than the marginal cost incurred to serve the marginal buyer of the original product. By shifting some consumers from copies to the original, increased copyright protection lowers the cost at which these 'shifted' consumers are served, increasing total welfare.[22] While surprising and interesting, the importance of this result should not be overstated: it depends on specific assumptions about the distribution of imitation costs and there is no evidence as to the likely magnitude of the effect. In particular, we would find it unwise to conclude from Novos and Waldman 's analysis that copyright protection can be increased from its current level at no social cost.

4. Trademarks

The economics of Trademarks is quite distinct from that of copyrights or patents. Trademarks, roughly speaking, are words, symbols or other signifiers used to distinguish a good or service produced by one firm from those produced by other firms.

The benefits of trademarks are several-fold. First, trademarks reduce consumer search cost by allowing consumers to quickly identify products with desirable attributes. This statement relies on several assumptions about the trademarked goods. First, for the trademark to have this benefit, it should be the case that the attributes of the product cannot be readily identified by simple inspection of the product. In other words, the trademarked good should be an experience good in the sense that a consumer must be able to consume it in order to evaluate its true characteristics. Second, the producer of the trademarked good must be able to maintain consistent characteristics in the product, including its quality, over time. Otherwise, past consumption would be no guide to future consumption of the same good. Consumers, in turn, would not be willing to pay more for a trademarked product because it would not reduce their search cost.

---

[21] See Takeyama (1994).
[22] Total welfare is the sum of consumer surplus and profits.

Trademarks also give firms an incentive to improve the quality of their product. Without an exclusive right on an identifying mark, a firm that is producing a lower-quality version of a good might be tempted to free ride on the firms producing high-quality versions by duplicating their trademarks and so misleading consumers into believing that the brands were equivalent. Since this would make it impossible for consumers to distinguish the high quality products, it would lower their willingness to pay for any product in this market. This would, in turn, lower the return to investing in quality and so would lower the incentive to create high quality products. Hence, the average quality of products in markets without trademark protection would be lower than in markets with trademark protection.

Clearly, other mechanisms are available to ensure product quality. For example, the legal system allows for damages to be paid to parties that have been subject to deceptive practices. Hence, a firm claiming high quality and, in fact, supplying low quality could be required to compensate its customers. The damage system has disadvantages, however, in the sense that excessive damages can create perverse incentives for customers to induce breach (such as sabotaging a jetliner in order to collect on a particular passenger's death) while inadequate damages might not have the required disciplining function on firms. Reputational costs, imposed through market mechanisms, can work alongside damages to ensure product quality without creating perverse incentives. For example, suppose that a trademarked product proved to have low, rather than high, quality. Consumers would impose a cost on the firm by refusing to purchase. These costs would not go into the pocket of the deceived customers, however, so that they would not have an incentive to induce breach. Furthermore, these reputational costs can be, and in fact have been measured to be, quite high[23]. This could serve to discipline firms. In fact, De Alessi and Staaf (1994) suggest that the reason why damages often are relatively low in deception cases is precisely because the market imposes discipline of its own.

The role of trademarks in ensuring quality can be likened to the reward effect that we have identified in earlier sections: legal protection of trademarks allows an

[23] See De Alessi and Staaf (1994) for a review of the empirical evidence.

investment in quality to be rewarded by repeat purchase and other reputation effects (such as word-of-mouth advertising). This reward is associated with some "monopoly power" over the distinctive trademark in the sense that others can be excluded from using the same or a confusingly similar trademark. Still, to the extent that identifying names (and, in particular, fanciful names) are potentially in infinite supply (and at low cost of development) this monopoly power is not associated with a static welfare loss, as it is not associated with exclusion of other identical products from the market. Further, the power to exclude under trademark protection does not extend to the functionality of the product[24]. Hence, it is not possible to exclude another firm from producing a physically identical product: it is simply impossible to identify it in a way that confuses consumers about its source. This means that there is no static welfare loss associated with monopoly power over a product's function under trademark law[25]. Notice that the diffusion of information effect is not present in the case of trademarks, as their *use* discloses all relevant information.

As a result, trademarks have mostly a positive incentive effect, which suggests that they should be legally protected as long as they are used. In fact, when trademarks are allowed to be protected without use, they can be stockpiled. There is some evidence that this stockpiling causes barriers to entry in some markets, as the field of potentially attractive trademarks is reduced[26]. In terms of the scope of protection, the economic benefits of trademarks are present as long as confusion is not present. Hence, the economically appropriate scope is one that permits marks as long as they are not confusingly similar[27].

Overall, then**,** the economics of trademark protection and the intellectual property law of those marks are broadly in line. There are, however, some issues involved in the *use* of trademarks that need to be discussed. In particular, umbrella branding

[24] See Landes and Posner (1987) for a discussion and examples of cases where functional elements, specifically shape, have not been permitted trademark protection.
[25] Trademarks clearly can contribute to the image of a product, which can serve to differentiate the product. To the extent that this creates variety, quality, or disseminates information about the product, it may create a static welfare gain. To the extent that a protected image creates a barrier to entry by raising the entry cost potentially borne by entrants, this could create a static welfare loss. As the economics of this role of advertising creates ambiguous recommendations, it will not be considered as a factor here.
[26] See Landes and Posner (1987) for more discussion.

could raise concerns about extension of monopoly power from one market to another and compulsory licensing of trademarks as a remedy could have welfare-decreasing consequences by reducing the incentive to maintain high quality. These will be discussed later in the paper.

Part II. IPRs and Competition Policy: An Economic Perspective

1. Introduction

In this section, we address two sets of issues. We first consider the *systemic design* of IP and competition Law, trying to identify their separate functions and goals and to clarify the nature of their interaction. Our basic message is that, although these two fields of the law have a joint impact on economic incentives and performance, there is little need for explicit coordination. In particular, we argue that the supposed 'conflicts' between IP Law and competition law can be resolved by abiding by a few simple rules. The second issue is how, given a systemic design, competition law should be implemented when monopoly power is based on intellectual property rights. We will argue that, as a general rule, the treatment of IP-based monopoly power should not differ from the treatment of monopoly power stemming from any other source. In practice, however, market power rooted in IPRs retains some distinctiveness as some practices (e.g. cross-licensing) are more likely to emerge. We give some examples of such practices and briefly discuss how they should be treated.

2. The Argument for Independence

Intellectual property law differs from competition law in both its function and its goals. Broadly speaking, the main function of IP law is to properly assign and defend property rights on assets that might have economic value. The main function of competition law is to regulate the *use* of (intellectual) property rights *when these rights are sources of market power*. This market power element is important as intellectual property law also regulates the *use* of the property rights that it assigns but

[27] This is, in fact, the legal scope. Details, and an application to the case of geographically remote users is discussed at length in Landes and Posner (1987).

without reference to monopoly power.[28] From an economic perspective, then, the main goal of intellectual property law should be to strike the right balance between the various effects identified in the first part of our paper. On the other hand, the main goal of competition law should be to minimise the adverse consequences of monopoly power.

It is also important to consider the fact that intellectual property law and competition law tend to intervene at different stages of the economic lifecycle of an asset. Property rights are generally assigned very soon after the asset has been created, while competition law only intervenes significantly later, once using the asset has become the basis for some market power. An important consequence of this difference in timing is that the information available when property rights are granted is not the same as the information available when competition law cases arise. In particular, competition law authorities are likely to have much better information about the economic importance of a given innovation and about the structure of the market(s) where the innovation is used.

The fact that the two fields of Law have distinct functions and objectives does not necessarily mean that they can be designed and implemented separately. In fact, there seems to be an unavoidable source of conflict between the two bodies of Law. While IP rights do not necessarily confer significant monopoly power, they can only be effective if they sometimes do. After all, it is the expectation of some monopoly rents that drive both the *reward* and *disclosure* effects discussed in part I. As competition Law ‘s effectively constrains an agent ‘s ability to exploit its monopoly power, the two approaches appear to be on a collision course. Moreover, as Competition Law tends to have access to more detailed information than was available at the time property rights was granted, there is a great temptation to revisit the trade-off between innovation incentives and the inefficiencies resulting from the use of exclusive property rights.

[28] Examples of (ab)uses controlled by IPR law includes practices that deceive consumers (e.g. deceptive pricing or advertising) and, more generally any form of competition that is not ‘on the merits’. Broadly understood, this could be (and has sometimes been) construed as including some abuses of monopoly power. In practice, we would argue that there is much to be gained by keeping issues of monopoly power out of IP courts, leaving then to be dealt with under competition law.

The conflict between IP Law and competition Law is less 'unavoidable' than it might seem. One reason for this is that the assignment of property rights handles the 'reward' and 'disclosure' effects discussed in part I by offering an *expected* reward to the rights-holder. In many cases, this promised reward will not in fact materialise, as the innovation fails to find a profitable market or is rapidly pre-empted by further advances. In other cases, though, the innovation might prove significantly more profitable for a longer period than was initially thought. When investing in innovation, inventors usually do not have a very good idea of where their efforts might lead them on this continuum from bad to good fortune. What matters, then, is the reward that they can reasonably expect to obtain *on average*.[29] This has two important consequences. Firstly, *(intellectual) property law can achieve its goals even if competition law limits the extent to which the rights-holder can benefit from the monopoly power that might be attached to her property rights*. This can be achieved by granting stronger (i.e. longer and/or broader) property rights initially if the ex post restrictions imposed by competition law are expected to be strong on average. The only requirement for this approach to succeed is that competition Law does not essentially expropriate every right that results in some market power. Secondly, *there is no need for (intellectual) property Law to react to small changes in competition Law*. In particular, individual case decisions are of no consequence except if they herald a forthcoming sea-change in the enforcement of competition law.

Overall, then, all that is needed is for IPR law to adjust slowly over time to perceived changes in competition Law. Faster adjustments are also possible. An example of this 'fast track' approach is the US Drug Price Competition and Patent Term Restoration Act of 1984 (also known as the Waxman-Hatch Act). The act extended the duration of the patents granted on compounds that have to go through lengthy FDA approval procedures. The protection granted to other types of intellectual property could be similarly strengthened if, for some reason, there was a significant tightening of competition Law in some sectors of activity.

[29] This argument is similar to the view taken in Ayres and Klemperer (1999) and in Shapiro (2003). Shapiro sees patents as conferring *partial property rights*. Shapiro elaborates as follows: "Nothing in the patent grant guarantees that the patent will be declared valid, or that the defendant in the patent suit will be found to have infringed. In other words, all real patents are *less strong* than the idealized patent grant usually imagined in economic theory" (P. 395). We simply extend this approach by adding that there is no guarantee that the right to exclude granted to a patent-holder, for example, will not be

Having discussed whether and how (intellectual) property law should accommodate changes in competition law, we now examine whether competition law should systematically revisit the trade-offs already considered in the design of (intellectual) property law. In particular, should arguments about the trade-off between 'static' and 'dynamic' efficiency be part and parcel of competition law cases – as they increasingly are -? The short answer to this question is no. The main reason for this is precisely that the relevant trade-offs have already been embedded in the design of the various property rights regimes. As we saw in the first part, these regimes appear to accommodate the essential differences between various types of assets rather well. Moreover, the specific rules that apply to different kinds of *intellectual* property closely resemble the socially optimal mechanisms that economic theory would recommend.

One could object that, even though property rights regimes optimally balance static and dynamic efficiency considerations, they do this based on the information available at the time property rights are granted. As we have just seen, competition authorities are likely to have better information at the time of their own involvement. Shouldn't this additional information be used to 're-optimise' and adjust the balance called for by the reward and disclosure effects? For example, we know that market structure can affect the private returns to investment. Why, then shouldn't we apply competition policy differentially across sectors in order to fine-tune the balance between investments incentives and efficiency losses? If, for example, we knew that less competitive markets increase private returns to investment (compared to their social return), then being tougher on firms operating in concentrated markets would in fact bring returns in such industries in line with returns in other more competitive sectors. The problem with this line of argument is that, in fact, *we do not know what the effect of market structure on investment incentives are*. Economic theory just does not have any robust prediction as to whether 'competition drives innovation' or 'invention is best nurtured – and financed – by large firms with significant monopoly power.[30]

---

limited *ex post* by competition law. What matters is the expected monopoly power associated with IPRs *ex ante*.

[30] For a summary of this debate, see Kamien and Schwartz (1982). Although not recent, this reference is not dated: the debate has not progressed significantly over the last two decades.

A further reason not to allow competition law judges to systematically revisit the trade-offs already considered by (intellectual) property law is what economists refer to as the risk of *regulatory opportunism*. As we saw in part I, once intellectual property is produced (and disclosed), the socially optimal allocation is for every economic agent to have free access to it. In other words, the optimal level of monopoly power *ex post* is none. As competition law only faces such ex post situations, there might be a strong temptation to limit the use of IP-based monopoly power so much that adequate rewards for investment in IP could no longer be provided. This temptation might even be stronger at the level of individual cases since, as we discussed above, a single case is unlikely to significantly the *expected* reward on which investors base their decisions. However, succumbing to this temptation would lead to a 'death of a thousand cuts', where the combined effect of apparently innocuous individual case decisions combine to wreck the delicate balance achieved by IP law.

The remedy against regulatory opportunism is commitment. In the case of competition law, commitment can only come from the clarity of how the law should be implemented. The clearer the rules, the stronger the commitment. It would therefore be advisable to explicitly state that competition law should respect the rights granted by (intellectual) property law and that the trade-off between static and dynamic efficiency is not a primary concern of competition law. This principle does not prevent the enforcement of (possibly strict) competition laws but it implies that conditions under which the use of monopoly power will be restricted must be as unambiguous as possible. The *essential facility* doctrine can be seen as a good example of such an approach. On the one hand, it is entirely consistent with the general respect of property rights that we advocate and acknowledges that property rights can only be effective if they do imply some monopoly power. On the other hand, it allows for a clear exception when the monopoly power associated with the property right is so large as to result in an unacceptable loss of welfare.[31]

[31] The economic logic behind this argument is similar to Gilbert and Shapiro 's argument discussed in part I: if increased monopoly power increases the deadweight loss faster than profits, then extreme monopoly power can be very damaging without necessarily adding much to the ex ante incentives to invest.

Before closing this section, we must point out that the arguments presented apply to investment in all types of economic assets, not just to intellectual property. This has a most important consequence: *as the distinct characteristics of various types of property are already adequately reflected in their specific property rights regimes, all types of assets should be treated equally by competition law*. From the point of view of competition law, the only relevant difference between assets is the degree of monopoly power that their ownership confers. For equal degrees of market power, further distinctions between asset is not only not required but counterproductive: by introducing artificial differences in the treatment of assets, competition law would only skew the relative returns that can be obtained from different types of investment and adversely affect he allocation of resources in the economy.

In particular then, claims that a firm with significant market power should be treated more benevolently because 'the source of its market power is intellectual property and being harsh would compromise innovation and the social benefits accruing from it' should be dismissed. The specificity of intellectual property, including the 'social benefits accruing from it' has already been taken into account in the special property rights regime that it enjoys.

The discussion in this section can be summarised in the following set of principles for competition policy:

a. Restraint. If competition law focuses narrowly on monopoly power, it risks dissipating the expected rewards that are essential to provide adequate investment incentives. This principle applies to both intellectual and 'real' property.

b. Each individual competition law decision might seem to have (and does have) only a small effect on expectations of reward but their combined effect can be devastating. There is therefore a need for a clear commitment not to revisit ex post the rights granted by IP law.

c. There is no need to systematically revisit the trade-off between incentives and economic efficiency in competition law cases.

d. Only large changes in ex post regulation (e.g. competition policy) calls for adjustment of property right regimes.

e. It is important to treat all sources of monopoly power similarly.

3. Applications and Specific issues

In this section, we discuss how the principles laid down in section 2 can be applied to patents, copyrights and trademark. As above, the main line of argument is that intellectual assets should be treated like any other source of market power. This, however, does not mean that the interface between intellectual property and competition law does not have any special characteristics. In particular, we will argue that some types of (potentially) abusive practices are more likely to arise in the presence of IP-based dominance. We will also see that the competition authority 's ability to detect abusive practices and to enforce adequate remedy can be affected by some of the special features of intellectual assets.

a. Patents

Intellectual property is only a concern for competition policy if it is the source of significant market power. This raises the question of how relevant markets – and a firm 's competitive position in these markets – should be determined. Do the traditional approaches to market definition, 'market power' and 'dominance' work well when intellectual property rights are involved? Subject to the caveats discussed below in the paragraph on mergers, the answer is yes. One should just be careful to identify *all* relevant markets. In the case of patents, one would generally expect both *upstream* and *downstream* markets to be involved. The downstream markets are all the product markets where the patented innovation can find a commercial use. The relevant upstream market is the market for 'knowledge' itself, where firms allow others to use their intellectual property through various forms of contractual arrangements.

In determining the relevant downstream markets, one should be mindful of the fact that there is no strict correspondence between the notion(s) of patent breadth and

the market power that the patent confers. An example of a 'broad' patent is Agracetus' EPO patent covering *any* possible type of transformation of soybean through genetic engineering techniques. If GM products had proved to be a commercial success in Europe, this patent would likely have resulted in significant monopoly power in the soybean market. On the other hand, consider another patent that would grant exclusivity to all genetic engineering methods of conferring glyphosate resistance to plants. This is also a rather broad patent as it covers all types of plants, wheat as well as cotton or fruit trees. On the other hand, as glyphosate is not the only high-performance herbicide available, plants that are resistant to other herbicides are still likely to limit the market power of the patent holder in any of the relevant downstream markets. In other words, the legal 'breadth' of the patent is 'spread' across a large number of relevant product markets (say one per plant of family of plants) so that the resulting market power needs not be a concern.

A. Mergers

In appraising a merger, competition authorities evaluate whether markets are likely to be significantly less competitive after the new entity is formed. This assessment involves both *unilateral* effects and *coordinated* effects. We will focus on the former. In practice, the first step of the competition authority 's analysis is to determine the likely effect of the mergers on *market shares* in the relevant market. These shares need not be based solely on the *sales* of the merging party. Often the parties ' share of industry-wide *productive capacity* is also considered because it is thought to be a good indicator or *potential* market shares.

As argued above, the same principles should be applied regardless of the source of potential monopoly power. In terms of mergers, this implies that, while shares in the relevant product markets are of course still relevant, one should also seriously scrutinise the merging partners ' intellectual property 'capacity'. As a first step, one should assess the firms' share in the upstream market for intellectual property. This involves a disclosure of all licensing agreements, i.e. of the actual 'sales' of IP. This however does not suffice: as for productive capacity, it is also necessary to assess the merging parties ' share of the existing stock of intellectual assets. This share should be computed for each of the downstream markets. This

presents two difficulties. Firstly, as discussed above, going from a specific patent to the markets that it can affect is not always straightforward. Secondly, it is not obvious how to weight the various patents in the firms' IP portfolio to obtain some aggregate measure of 'capacity'.

At the level of the EU at least, publicly available information and our own experience suggests that a thorough appraisal of merging parties ' IP positions rarely takes place. Even considering the difficulties of measurement involved, we would argue that more rigorous assessments are both doable and highly desirable. By making IP-intensive mergers more likely to be approved that mergers involving other sources of market power, the current practice might artificially bias the allocation of resources toward IP-incentive sectors.

B. Licensing

In most jurisdictions, patent licensing is the object of a host of special competition law rules.[32] From the point of view of economic analysis, this is rather confusing as there is essentially no reason to treat patent licensing agreements differently from any other kind of vertical contract. Hence, in the EU for example, most aspects of licensing contracts can perfectly well – and should – be assessed according to the Commission 's 'vertical guidelines'. Accordingly, special dispensations, such as the EU 's block exemption on patent licensing, do not make much sense. In fact, by treating patent-based monopoly power differently (and-arguably – more leniently -) than market power based on other types of assets, such a policy distorts the economy-wide allocation of resources (see part I).

One caveat to this general principle is that, because of the nature of the 'input' being sold, licensing contracts must often include clauses aimed at safeguarding the integrity of the innovation.[33] In other words, some contractual clauses might be indispensable as, without them, the value of the intellectual property right might be lost. Clearly such clauses can only justify a differential treatment of IP-based market power if they would not be necessary to protect the value of other types of

[32] See Anderman (1998) for the EU and both the old 'No Nos' of licensing and the more recent 1995 DOJ/FTC *Antitrust Guidelines for the Licensing of Intellectual Property* in the US.
[33] See Anderman (1998), pp. 102 – 129.

contractible assets. The main source for such a discrepancy is the public good nature of intellectual property, i.e. the fact that it is an 'input' in the production process that is not effectively destroyed when used. This has several implications. The first issue is that of *resale*. A manufacturer purchasing ball bearings from a firm with market power cannot resell these inputs to another downstream customer without depriving itself of its use. By contrast, since knowledge is non-rival, a licensee can easily resell the knowledge acquired from the licensor to a third party without restricting its own use of the technology. In fact, if the licensee and the firm it resells the technology to operate in distinct market, the resale does not impose any cost at all on the licensee. Accordingly, it is perfectly legitimate to allow clauses that forbid the divulgation of the licensor 's intellectual property to third parties, at least until that property has fallen into the public domain. Such clauses could include, for example, a ban on sub-licensing or even assignment.

A second issue, is that of the reputation *of the technology*. Even if licensees do not use the licensor ' s trademark, their behaviour can damage the value of her intellectual property. Consider for example a new contraceptive device. Even if faulty products are clearly identified as being made by a given licensee, they can damage consumer confidence in the new contraceptive method itself. This would damage the profitability of all licensees and, therefore decrease the revenues of the licensor. In fact, when consumers are initially quite uncertain about the new technology, the behaviour of a few rogue licensees might affect the technology very viability. One might therefore want to treat clauses aimed at ensuring some quality control (e.g. prohibition to deviate from the technology or prescribed production methods) rather leniently.

A somewhat different problem is that it is difficult to measure the 'intensity' with which the licensee uses the contracted input. The issue here is not that this difficulty might 'destroy' the value of the patent but, rather that it might make patent-based monopoly inherently less profitable than other types of monopolies. When dealing with a physical input, the intensity of used can be assessed directly from the number of units purchased by the 'downstream' firm. The upstream firm can then easily vary the price of the input depending on the total quantities purchased. In order to duplicate such contracts, and therefore be able to 'extract surplus' as efficiently as

the seller of physical inputs, a licensor must be able to condition the payments received from the licensee on some other measure of 'intensity', such as the relevant sales of the licensee. This can justify stricter clauses – such as some control over distribution channels downstream, aimed at making such measure reliable.

While patent-based monopoly power should not receive differential treatment, it gives rise to potentially abusive conducts that are rarely observed when market power comes from the ownership of other types of exclusive assets. The competitive implications of these types of conduct – not some broad theoretical 'conflict' between the two bodies of Law- are the proper subject matter for those interested in the interface between IP Law and Competition Law[34].

1. Cross-licensing

From the point of view of competition policy, cross-licensing and patent pools are probably the most significant IP-specific practices. Even in this case, though, one can think of non-IP equivalent. For example, a contract whereby two airlines which have a stranglehold on two different airports agree to give each other access to their gates, terminals and/or other ground facilities looks very much like an agreement to share various technologies. Still, patent pools and cross-licensing are observed much more often than their non-IP analogies. We will initially focus on cross-licensing, leaving some of the specific features of patent pools for the end of this section.

In assessing the antitrust implications of cross-licensing, it is vital to determine whether the technologies involved are (broadly) *substitutes* or *complements*. Technologies are *substitutes* if they (potentially) compete with each other. This does not necessarily mean that the scientific principles on which the two technologies rely must be similar. For example, two patents on separate pain relievers are substitute even though the chemical compounds and the physiological mechanisms involved might be very different. On the other hand, technologies are complement if using them jointly enables a firm to improve the quality of its product and/or lower its cost of production. One might believe, for example, that combining the DVD patents held by a number of firms would allow each of the firms involved to present more

attractive products to consumers. In fact, in cases where different parties hold 'blocking' patents on different aspects of a technology, cross-licensing might be the only way to ensure that the new technology is used at all. Because of this 'value-increasing' or 'cost-decreasing' effect, cross-licensing of complementary technologies should, as a rule, be given the benefit of the doubt: in the absence of some specific, documented, competitive concern, they should escape antitrust scrutiny. Quite the opposite principle applies to the case of substitute technologies: given the lack of any obvious benefit, they should generally be considered with suspicion.

There are two main reasons to be wary of cross-licensing agreements between firms. The first issue relates to the structure of the licensing payments. As Katz and Shapiro (1985) and Fershtman and Kamien (1994) have shown, competing firms can replicate the monopoly outcome by choosing appropriate levels of royalties. The intuition behind this result is relatively straightforward. Consider cross-licensing between two firms, A and B that compete in the same market and assume that each firm actually uses the technology that it licenses from the other. Assume further that the royalty payment is linked to the volume or value of the sales made using the licensed technology. When deciding how hard to compete (i.e. how much to produce or what price to set), firm A considers two factors that would not be present without cross-licensing. Firstly, A must now pay a royalty to firm B. This effectively raises firm A's cost of production, leading to less aggressive behaviour.[35] Secondly, firm A also considers the effect of its behaviour on the flow of royalties that it is getting from firm B. As more aggressive behaviour on A 's part reduces both the output and the profitability of its rival, it also decreases A 's licensing income, leading it to adopt a less aggressive stance. Of course, the same reasoning also applies to B. The end result is that both firms compete less harshly, moving the industry closer to the monopoly outcome.

Notice that this *collusive* effect does not require any explicit or even tacit coordination of the two firms 'actions: the cross-licensing agreement simply modifies the firms ' incentives to ensure that the uncoordinated equilibrium is less competitive

[34] For an alternative, and complementary view of this, see Scotchmer and Maurer (2002)

than before. In fact, this is very similar to the effect of cross-ownership of shares, to which several antitrust authorities are now paying increasing attention: a firm owning shares in a rival will compete less intensively since competition hurts its rival and, therefore, the value of the firm 's shareholdings in its rival. As a matter of consistency then, it would make little sense to scrutinize cross-share holdings and ignore cross-licensing agreements. It is also worth noting that the collusive effect of royalty payments arises irrespective of whether the technologies licensed are substitutes or complements. In the later case, then, one would have to weight the potential benefits of cross-licensing (see above) against the resulting decrease in the intensity of competition. As a final remark, notice that the effect discussed above depends crucially on the *structure* of the royalty payments. Fixed payments, i.e. payments that are not linked in any way to the performance of the rival, would not affect the competitiveness of the industry. This suggests a possible remedy that would remove any ambiguity for the case of complementary technologies: if the parties are willing to rely mostly on lump sum licensing fees, then cross-licensing of such technologies would not raise any serious antitrust concern.

Cross-licensing agreements can also restrict competition through a very different channel by acting as *facilitating practices*, i.e. practices that facilitate *tacit collusion* between rivals.[36] Tacit collusion refers to an implicit agreement to keep prices high (or quantities low). They are enforced through implicit threats: as long as other firms hold their side of the bargain, everybody else does too. But as soon as one firm 'deviates' by setting a lower price or producing more then the industry gets into a 'punishment' phase, where all members of the tacit agreement set much lower prices (or sell larger quantities). It is the threat of these costly punishment episodes that ensures that the firms prefer to abide by the implicit agreement. The more effective this threat, the more collusive (i.e. closer to the monopoly outcome) the industry equilibrium that can be sustained. It is precisely by making the punishment harsher that cross-licensing agreements can be anti-competitive. To explain how this comes about, it is worth considering a number of distinct situations.

[35] This is straightforward for the case of *per unit* royalties as they directly increase marginal cost. If the royalty is on the value of sale, they reduce the firm 's marginal revenue while leaving its marginal cost unchanged. This too leads the firm to behave less aggressively.

Assume first that the two firms, A and B *do not compete in the same market*. For example, A might have a patent for an antidepressant while B might have a patent for a drug that fights stomach ulcers. In the absence of cross-licensing, each firm might be tempted to develop its own drug in order to enter the other firm 's market. The firms could try to keep their rival out of their market by threatening to invade the rival 's own turf if and only if the rival invades first. Unfortunately, this threat is unlikely to be very effective: if A does in fact move into the other firm 's market, it will take time for its rival to react by developing an antidepressant of its own. During that time, A enjoys monopoly profits I antidepressants *and* gets a share of the anti-ulcer market. If B 's reaction time is long enough, this prospect would prove too attractive and both firms would in fact get onto each other 's turf, increasing competition. Suppose now that the two firms have a cross-licensing agreement. This means that each firm could now react much more quickly if its rival decided to breach the implicit agreement. Anticipating this quicker reaction, both firms find it more profitable to stay in their own market and respect the monopoly of the other firm. This anti-competitive effect of cross-licensing has two notable features. Firstly, it applies to firms which *do not compete in the same market*. Secondly, *the firms do not actually use the technologies that they obtain from the other firm*. This is an important feature and strongly suggests that, when tacit collusion is a concern, the actual use of cross-licensed technologies should be closely monitored.

Now assume that the two firms actually compete in the same market, say the market for pain relievers. Their products are substitutes, but not perfectly so. For example, firm A 's product might be aspirin-based while B 's relies on paracetamol. The two firms might already be able to support some level of tacit collusion even in the absence of cross-licensing: firm B knows that if it lowers its price, A will retaliate by starting a price war. As the products are substitutes, A 's retaliation would hurt B significantly. However, as before, the retaliation would be even more effective if A could retaliate by selling a product that is an even closer substitute to B 's product. This is precisely, what cross-licensing would make possible. As firms would fear retaliation more, they would be able to further decrease the intensity of competition in

[36] See Eswaran (1994) for a formal analysis.

the market. In this context, the firms do compete in the same market but, as above, they do not actually use the technology obtained from their rivals.[37]

Overall then, cross-licensing is likely to facilitate tacit collusion irrespective of whether they involve competing technologies or technologies that find their applications in distinct market. The fact that the firms do not broadly use the licensed technologies should be seen as prima facie evidence for this collusive effect. Importantly, this concern does not apply to the cross-licensing of complementary technologies.[38]

2. Patent Pools

From the point of view of economic analysis, patent pools are very similar to cross-licensing: a number of firms gives each other access to a number of their patents. The payments schemes involved vary widely. Some pools grant free access to all members, others involve elaborate royalty schemes. In this respect, as discussed above, the main concern would be a systematic reliance on output or sales-related royalties. Like cross-licensing, patent pools also facilitate tacit collusion. There are, if anything, even more suspicious in this respect, for two reasons. Firstly, given that pools typically include a large number of patents, one would not expect all members to use all technologies anyway. This makes it harder to distinguish 'innocuous' pools from those meant to reduce competition between members. Secondly, the number of patents involved also multiplies the potential for 'multi-market contact' between members pool member, making tacit agreements even easier to support.

A distinct feature of patent pools is the conditions of access that they set for *non-members*. The potential for abuse is obvious. For example, refusal to grant

[37] Theoretically at least, cross-licensing can facilitate tacit collusion even if the firms actually use the technologies that hey acquire from other. This is especially likely in a situation where technologies apply to different markets. As Bernheim and Whinston (199x) have shown, under some conditions, extending he number of markets where the same firms compete increase their ability to support collusive outcomes.

[38] It *migh*t apply if cross-licensing the two complementary technologies help make the products of he two firms more similar. This, however, is a much less robust effect than in the case of independent or substitutable technologies.

access to third parties on terms resembling those available to members would, if the pool members have sufficient monopoly power, amount to collective foreclosure. There would in this case be absolutely no difference between a patent pool and discriminatory access arrangements among a group of powerful airlines. Another common practice has more subtle antitrust implications. Often, third party are only offered access *the whole set* of patents in the pool, i.e. they are charged a single price for what is effectively a 'bundle' of intellectual property rights. Such tactics could be used to *leverage* the monopoly power of that pool members enjoy in one market into another. The precise economic argument can be found in Whinston (1990).[39] Its basic flavour can be obtained from the following example. A firm (or a patent pool) has patents on two types of vaccines, one against polio, the other against German measles. Assume further that the firm ' s polio vaccine is the only one available, i.e. the firm is a monopolist in the market for polio vaccines. On the other hand, it faces (potential) competition in the market for the other vaccine. The firm licences its patents to companies that actually make the vaccine. Suppose that instead of licensing he vaccine separately, the firm *only* offers the two licenses jointly. What this means is that, in order to realise the profits corresponding to its polio monopoly, the firm now *must* also induce its customers to buy its German measles vaccine. This makes the firm much more aggressive in the market for German measles vaccines than it would have been, had it decided to license its two technologies separately. In particular, the firm might be willing to sell its German measles vaccine at an *implicit* price[40] that is below the cost of its main (potential) rivals. The prospect of such fierce competition would discourage potential entrants and might even induce existing competitors to exit. If this occurs, the firm will have successfully used its monopoly power in the polio market to enhance its position in the market for its other vaccine.

A subtlety of this argument is that it does not apply when the two patents involved are strict complements. Suppose that firm A had a patent-based monopoly on a medical diagnostic machine that uses films but that various types of films could be used on the machine. Because the machine cannot be used without films and vice-

[39] This argument was amply discussed in the context of the Microsoft cases, both in the US and in the EU.

[40] Suppose that the royalty that the firm would charge for the polio vaccine, were the patents to be licensed separately were r^M and that the royalty that it charges for the two licenses jointly is r^B. The implicit royalty charged for the German measles vaccine as part of the bundle is simply $r^B - r^M$.

versa, the firm could extent its monopoly in the 'machine' market into the market for film without selling machine and films jointly. Since the customers must use the two products together anyway, the firm could charge a very low price for its film, driving out the competition, and simply recoup this by selling the machine at a price in excess of its stand-alone monopoly price.[41] This is therefore a further reason for looking less favourably on pooling of patents that are not strong complements.

Overall, then, there are overwhelming reasons to be suspicious of patent pools. There is a strong argument for requiring that all patent pool agreements be notified to the antitrust authority, if not for banning them outright.

3. Settlements[42]

Patent pools and cross-licensing agreements often arise as part of litigation settlements. In such cases, their welfare properties cannot be properly assessed without considering the potential benefits and costs of having firms reach such a settlement rather than proceed with litigation. Moreover, this also means that agreements to settle patent litigation should not be the sole province of patent Law and IP courts: they should be subjected to antitrust scrutiny. Since more than 95% of US patent litigation cases are settled (see Lanjouw and Schankerman, 2003), this issue is of more than academic interest.

Clearly, settlements are only reached if they are privately beneficial to each of the contracting parties. Social costs and benefits are those affecting parties that are not involved in the agreement, mostly consumers, the government and – possibly – other firms. On the benefit side, avoiding litigation saves on direct court costs and helps relieve court congestion. On the cost side, are the fact that settling a case might prevent the establishment of a socially useful precedent on a point of Law and any potential loss of competition. The resulting decrease in competition can take three forms. Firstly, market rivalry between parties would presumably have persisted over the period of litigation. Secondly, there is a chance that the outcome of continued

[41] Strictly speaking, this argument only applies exactly when the two complementary products – or license- must be consumed in fixed proportions. It is only in that case that the traditional Chicago-school- claim that foreclosure cannot occur 'because there is only one monopoly profit' applies with full force.
[42] This section is based on Shapiro (2003)

litigation would have been to invalidate or seriously limit the contested patent, eliminating or reducing the monopoly power of the patent-holder. Finally, as discussed in the previous section, the settlement itself might support more collusive behaviour in the post-settlement market.

Given these costs and benefits, how can we distinguish between 'collusive' and 'pro-competitive' settlements. Shapiro (2003) proposes to rely a *principle of consumer neutrality*: an agreement will be deemed to be pro-competitive if it leaves consumers at least as well off as they would have been had the parties seen the litigation to its bitter end. Shapiro 's core result is that, under quite general conditions, there always exists a settlement that makes the litigating parties better off without hurting consumers. In other words, respecting the principle of consumer neutrality should not prevent litigating parties from reaching a settlement, it only imposes restrictions on the *types* of settlements that might be reached. In practice, however, the information required to determine which agreements pass this test might be hard to obtain. In particular, since one must compare the outcome of the agreement to what would have occurred, had litigation proceeded, one needs to assess the strength of the contested patent and this must be done without the benefit of a full trial on the issue.

4. Grant-backs

As part of the conditions for licensing their technology, many companies require their licensees to 'grant back' to them any improvements that they make. The precise agreement can take a variety of forms. The grant-back can be free or involve some payments from the original licensor and the licensor might or might not enjoy the exclusive benefit of the improvements. The grant-back can also be *unilateral*, in which case improvements only flow from the licensee to the licensor or *mutual*, in which case the licensee also receives the further improvements discovered by the licensor.

To discuss the economics of grant-back, it is useful to take EU competition Law as a point of reference. In a nutshell[43] under article 85(1), the Commission has

[43] For more details, see Anderman (1998), pp. 109 – 118.

no objection to grant-back clauses that are both *non-exclusive* and *mutual*. The Commission further distinguishes between *severable* improvements, which are those that can be used independently of the original licensed technology and *non-severable* improvements which can only be exploited jointly with the original licensed technology. The Commission 's position on severable improvements is strong: the licensee has the right to use and license this improvement both during and after the term of the initial licensing contract. The Commission also takes a dim view of any attempt by the original licensor to obtain rights over the severable improvement *beyond the term of the original licensing agreement* without an appropriate quid pro quo. What constitutes an "appropriate" quid pro quo is however not clear. There seems to be some preference for continuing reciprocal exchange of the right to use each other 's technology over the simple payment of royalties by the former licensor. The treatment of non-severable improvements differ mostly during the term of the initial licensing contract, as an *exclusive* – but still mutual – Licensing of improvements would not be deemed to raise significant competitive concerns.

The most obvious effect of grant-back clauses is that they tend to decrease the parties ' incentives to invest resources in seeking to improve the technology. To see this, consider the simple case where the licensee must grant back its improvements to the licensor for free. Compared to a situation where the licensee could negotiate *ex post* a reward for transferring its new know-how, the grant-back clause decreases the returns that the licensee can obtain, discouraging investment. Requiring that grant-back clauses ensure the mutual exchange of improvement does not help. On the contrary, this also discourages innovation as it decreases the licensor 's own incentive to improve the technology. Also– by guaranteeing that the licensee gets some technology improvements anyway *mutual* grant- back clauses might[44] also dull the licensee 's own incentives to innovate even further. The Commission 's preference for reciprocal exchange is not therefore particularly well founded.

One should add that, in terms of economic analysis, the idea that 'less innovation is bad' is not particularly compelling either. While consumers are likely to benefit from more innovation, society as a whole might not as the firms might be

[44] This second effect can actually go either way: getting better technology from its rival might induce lesser effort from the licensee or might, on the contrary both increase the licensee 's incentives to forge ahead and its ability to do so. See Choi (2002) pp. 817 – 818.

investing too heavily in R&D. The reason for this goes back to the *business stealing* effect discuss at the beginning of this paper: while some of a firm 's reward from innovation comes from the greater social value that it creates, another part comes from the profits that it diverts from other firms. The first part is socially useful, the second part is not.[45]

The disincentive effect of grant-back clauses could of course be avoided if each party agreed to make an appropriate payment for the new know-how that it receives. In that sense, the Commission 's preference for a 'quid pro quo' appears to be justified. However, in order to be effective, such a quid pro quo must be *conditional*, i.e. the contract should be such that each party only receives something from the other *if it does indeed produce improvements and make them available to the other party*. Hence, a broad agreement to exchange all improvements without payment would not help at all since a party would receive the other party ' s improvements regardless of whether or not it comes up with improvements of its own. Another issue is that conditional payments are hard to determine ex ante. By definition, future improvements cannot be described – and therefore valued – accurately before they have been obtained. This makes the inclusion of pre-set tariffs in a grant-back clause unlikely. The alternative, then, is simply to negotiate the price of the transfer ex post, i.e. once an improvement has actually been made. This approach seems quite feasible in the case of severable improvements: as the number of potential users is not limited by the availability of the original technology, the licensee should be able to obtain a reasonable return on its investment. If the improvement were non-severable, then the parties would run into a traditional *hold up* problem. This is clearest in the case where the licensor only has one licensee. Since the two parties are the only one who can actually use the non-severable

[45] Van Dijk (2000) develops a model where the firms ' private incentives to innovate are too large. In such a case, grant-back clauses might be desirable since they help scale R&D back. Choi (2002) also points the possibility that grant-backs can reduce socially excessive R&D. one could however take the following pragmatic attitude. Whether or not R&D is indeed insufficient, public policy makers certainly act as if they believed that it is: measures aimed at *increasing* innovation are many and they receive a great deal of publicity. On the other hand, examples of measures that are publicly advertised as aiming at reducing innovation are rare (in fact, we do not know of any). As a matter of *revealed preference*, then, one might be justified to assume that privately financed innovation tends in fact to fall short of what would be socially desirable.

[45] This second effect can actually go either way: getting better technology from its rival might induce lesser effort from the licensee or might, on the contrary both increase the licensee 's incentives to forge ahead and its ability to do so. See Choi (2002) pp. 817 – 818.

improvements, these innovations have no value outside of the specific licensor-licensee relationship. This makes it easy for the 'buying' party to obtain the improvement at a price much below its actual value. This in turn means that the two parties would have insufficient incentives to pursue non-severable improvements.

Overall, then, economic analysis appears to provide little support for the Commission 's preference for the mutual exchange of improvements. Furthermore, the arguments presented suggest that there is no benefit to agreeing *ex ante* to the exchange of severable improvements. On the other hand agreeing on ex ante conditional payments for non-severable improvements might help resolve a hold up problem, increasing the firms ' incentives to innovate.

The issue of exclusivity is relatively straightforward. There is absolutely no reason to let the licensor obtain through the grant-back more 'exclusivity' than already conferred by her patent on the initial technology. One implication of this principle is that the licensor should never be allowed to demand *ex ante* that the grant-back of *severable* improvements be exclusive.[46] The implications for *non-severable* improvements are rather different. During the validity of the initial licensing agreement, the licensor already has the power to prevent any party that is not authorised to use the main technology from using the improvement. A clause requiring exclusivity for the licensor, *and all of its licensees* would therefore be redundant and, as such, would not further damage competition. On the other end, one should take a dim view of exclusivity clauses that prevent the licensee from using its own improvement or from making them available (at a price) to all licensees of the original licensor.

The previous arguments assume that the original technology has already been licensed. Choi (2002) examines the effect of grant-back clauses on the original inventor 's incentive to license in the first place. Choi 's model has several important features. There are two firms that are involved in a repeated innovation race. These two firms only compete in the market for technologies, i.e. they sell their technology to the same potential customers. The two firms do not compete in the same product market(s). Initially, one of the two firms has a technology that it considers licensing

to its rival. The benefit from licensing is the revenues that it generates. The cost is that, by licensing its technology, the licensor essentially gives its rival a leg up in the race for further innovations. Still Choi shows that *if the two firms can write complete contracts* then the licensor always sells its highest quality technology to its rival. In practice, however, parties are unlikely to be able to write complete contracts. Choi assumes that the initial licensing contract between the two firms involves *moral hazard.* More precisely, the *quality* of the technology licensed is not contractible: while the two firms may be able to observe this quality, a third-party could not make any contractual clause based on observed quality unenforceable. If payments cannot be linked explicitly to quality, how then can the licensor ensure that it gets higher revenues from licensing a higher quality technology? The answer lies in royalties that are tied to the output of the licensee: as better technology will result in greater sales, the licensor 's revenues are tied indirectly to the quality of the licensed technology. Unfortunately, royalties are imperfect instruments: they do not extract the full value added by the licensed technology. There will therefore be situations where the revenues actually obtained through per unit royalties do not cover the cost from increasing the rival 's ability to compete, even though the true value of the technology transfers would exceed this cost. In such cases, the licensor would not license its best technology even though this is the socially optimal thing to do. Choi 's key result is that, in the presence of this moral hazard problem, including grant-back clauses in the licensing contract can ensure that the best technology is always licensed.

Choi 's analysis can therefore be seem as providing a rationale for a more lenient antitrust treatment of grant-back clauses. However, one should consider the following caveats. Firstly, while grant-back clauses ensure that the best technology gets licensed in the first place, it also affects the rest of the repeated R&D race between the two firms. These effects, as Choi admits, have ambiguous welfare consequences. Secondly, the model is a little too favourable to grant-back clauses as it assumes that they can be enforced without facing a *moral hazard* problem of their own. This seems extreme as, in practice, enforcing grant-back clauses also requires that the parties make sure that they are getting 'the best' improvement obtained by

[46] Licensee and licensor should of course be free – subject to the usual antitrust limits – to voluntarily reach an exclusive agreement ex post.

their rival. If the quality of the initial technology could not be contracted upon, why should we assume that the quality of improvements can be?

5. Compulsory Licensing

i. Refusal to License

As we have argued, a patent provides temporary property rights on an intellectual asset. This asset might prove useful in the production of a number of goods and services to which correspond various downstream markets. While the patent confers exclusive property rights it does not necessarily confers monopoly power. If it does then, subject to the caveats discussed above, it should be treated like any other source of market power. In particular, any refusal to grant access to the protected intellectual asset should be assessed in a manner consistent with the competition authority 's policy on vertical restraints. In the EU, for example, this would mean that a refusal to license the intellectual property to a (potential) downstream competitor would be deemed unlawful if the patent holder has significant market power in both the upstream and downstream markets. In other words, licensing could be made compulsory even if the patent is not absolutely essential to compete in the downstream market. The correct test in the upstream market is whether the number, type and ownership of alternative technologies is such that the patent-holder enjoys significant market power. Hence the exclusive nature of the property rights granted by a patent should not automatically translate into a right to use the patent exclusively.

The issue of compulsory licensing also arises in a rather different context where another firm comes up with an innovation that cannot be used without access to the original patent. This can be because the production of a commercially viable product requires that both innovations be combined or because the second innovation infringes the original patent. This situation is different from the 'vertical restraint' case considered in the previous paragraph because the potential licensee (potentially) competes with the patent owner in the *upstream* market, i.e. in the market for innovations. It is useful to distinguish between two scenarios, one where the two innovations are complements and one where they are substitutes.

The original patent and the new innovation are complements if both are needed in order to serve the relevant downstream market(s). In the field of genetically modified crops, for example, one might need to combine a patent on a specific DNA sequence (and its use) with a patent on a manner of introducing the DNA in a cell in order to produce a genetically modified seed. From an economic perspective, this does not give rise to any specific competition policy concern[47]. In particular, one would not expect the original patent holder to refuse to enter into some form of licensing agreement with the new innovator: since the patent holder could not enter the relevant downstream markets on its own anyway, any kind of licensing deal must be better than none. Still, such theoretical certainty is cold comfort if a refusal to license is actually observed. If this were the case, it is important to try to understand what motivates the refusal. A possible explanation is that the technologies are in fact not complementary. In practice, telling complements from substitutes is not necessarily easy. In the recent EU Microsoft case, for example, operating systems for PCs and operating systems for servers might have been construed as complements but the Commission argued that, in a dynamic perspective, they were actually substitutes.

Let us now consider a situation where the new innovation represents an (infringing) improvement on the original patent. As in the case of complementary technologies, the improvement represents an additional source of 'value' that the original patent-holder should be eager to share through some form of licensing. However, *unless the licensing agreement can be designed to effectively enforce collusion between the two firms*, allowing the new invention to be introduced also increases competition in the downstream market(s).[48] If this pro-competitive effect dominates, then the patent-holder would refuse to license even though licensing would be socially desirable *ex post*. Since competition policy will (and should) often frown on collusive licensing agreements, refusals to license are likely to be rather frequent. Does it mean that compulsory licensing should be imposed? As we will see, the answer depends on the terms of the compulsory license.

[47] The 'vertical restraint' concerns discussed above would of course remain. If the two innovators jointly have significant market power in the upstream market, they might still be required to license both of their technologies to some downstream competitors.

ii. Conditions of the compulsory license

As we have just seen, compulsory licensing is most likely to be called for in two types of situations: when the patent holder has significant market power in both the upstream and downstream market and when another firm develops an infringing innovation. However these two cases call for different principles when it comes to setting the terms of the licensing contract..

In the infringing innovation case, the object of compulsory licensing is to ensure that a socially useful innovation actually gets introduced, not to increase competition downstream. This objective can be achieved by choosing the royalty according to Baumol and Willig 's ECPR formula. This formula sets the 'access charge' paid by the licensee is equal to the marginal cost of granting access (likely to be close to zero in the case of IPR) plus an amount reflecting the profits lost by the licensor because of increased competition downstream. In other words, the license contract would be such that the original patent holder is made (at least) as well off as if it refused to grant a license. License terms that are less favourable to the licensor would amount to an ex post revision of the scope of the property right initially granted. As we have discussed above, such revisions are undesirable as they undermine the implicit 'contract between society and innovators.[49]

In the 'vertical restraint' scenario, the point of compulsory licensing is to ensure sufficient competition in the downstream markets. IPRs are, in this respect, treated like any other source of monopoly power. The terms of the compulsory licensing agreement can be determined in two manners. In the first approach, the competition authority would undertake to set the level of royalty, assuming therefore a quasi-regulatory role. Since the main concern is to promote downstream

---

[48] See Scotchmer and Green (1990) for a similar idea.

[49] Of course, such ex post revisions are fully justified if they are aimed at preventing abuse of monopoly power, as in our 'vertical restraint' case. Not doing so would amount to treating IPRs more favourably than other sources of monopoly power , which, as we have already argued, would distort the allocation of resources.

competition, the level of royalty should not be too high. In particular, it should be significantly below the level determined by the Baumol – Willig ECPR formula. The drawback of this system is that it can only work well if the competition authority has quite accurate information about the industry. Alternatively, the competition authority could simply decide how many downstream competitors are needed in order to insure sufficient competition. The rights to obtain one of the licenses can then be auctioned to the bidders offering the highest levels of compensation to the paten-holder.[50] Of course, the competition authority should make clear that only agreements that would pass traditional antitrust scrutiny (as outlined in paragraphs 1 through 4 above) are permissible.

6. Extending the legal patent 'monopoly' beyond the life of the patent

Several practices raise the question of whether the patent-holder is effectively trying to extend its monopoly power beyond the length of the patent. Examples include licensing contracts that require grant-backs or payments even after the licensor legal protection has expired. Such practices are surely an issue at the level of intellectual property law since they might be seen as violating the implicit contract whereby inventors receive protection for a pre-determined period of time as reward for their innovative efforts and as incentives to divulge economically useful information (see part I). But should these practices raise antitrust concerns?

We will limit ourselves to the case of post-patent payments. On the positive side, this can be seen as deferred payments that reduce the immediate financial burden of the licensee. In other words, the practice is equivalent to a loan from the licensor to the licensee that is paid back in instalments after the patent (or the licensing agreement) expires. Of course, this raises the question of why the licensee should receive this 'loan' from the licensor rather than from some standard financial institution. A possible reason is that the licensor has privileged information about the quality of the technology licensed. A bank might be reluctant to finance the acquisition of a technology of uncertain quality. The licensor should have no such qualm.

[50] To avoid the progressive disappearance of downstream competitors as some of them might fail, one might want to insist that these licensing agreements be potentially transferable to an alternative

To qualify as pure deferred payments, the fees paid after the expiry of the patent should be independent of the licensee 's post-patent use of the technology, i.e. they should either be pre-set lump sum payments or they should be linked to the licensee 's use of the technology while the patent was valid. Royalties that are linked to the output of the licensee in the post-patent period are potentially more problematic since they make the licensee less competitive even after the patent has expired. If the licensor competes in the same product market as the licensee, these continuing royalties can be seen as a way to raise the costs of a rival. One should however refrain from concluding that such royalties are undesirable. After all, such royalty schemes must be agreed by the licensee. Since the licensee knows that it will be legally able to access the technology for free once the patent has expired, it will never agree to a total payment that exceeds its willingness to pay in order to get access to the technology *while the patent is valid*. The relevant question then is whether, *for a given total (discounted) payment*, lower per unit royalties over a longer period are preferable to higher per unit royalties that stop at the end of the patent 's life. This problem is remarkably similar to Gilbert and Shapiro 's comparison of patent 's breadth and length. In fact, we already mentioned in part I that the size of the per unit royalty that can be charged was one possible example of patent 'breadth'. Applying Gilbert and Shapiro 's result, we can therefore conclude that deferred per unit period will actually improve welfare if the per period deadweight loss increases more quickly than the licensor 's per period revenues as the rate of per unit royalty increases.[51] If it does not then deferred per unit royalties are undesirable. In practice then, there seems to be little cause for systematically looking at such extended royalty scheme with alarm.

A further concern might be that the patent holder is leveraging its monopoly power during the patent period to increase its monopoly power in the post-patent period. This claim would rely on the foreclosure argument of Whinston (1990), which we already discussed in the section on patent pools. In this case, the 'monopoly market' would be a relevant product market while the patent is valid while the market

---

licensee.

to be foreclosed would be the same market after the patent has expired. By asking for payments in both periods, the licensor would essentially be 'tying' the sale of its technology in both markets. This analogy is in fact incorrect. In Whinston, tying works because it eliminates actual or potential competitors from the potentially competitive market. Here, however, there is one fundamental source of 'competition' that cannot be discarded: the fact that, in the post-patent period, the licensee can simply use the technology for free. In other words since the licensee does not need to obtain the technology from any other 'supplier' in the post-patent 'market' there is no room for the licensor to artificially expand its patent-based monopoly power.

b. Copyrights

Many of the arguments presented in the section on patents also apply to copyrights. We will therefore be brief and concentrate on a few distinctive aspects of copyrighted intellectual property as a source of monopoly power.

A significant difference is that individual copyrighted material is rarely the source of significant monopoly power. This is a direct consequence of the smaller breadth of copyright protection. From genetically-modified crops to Laser technology or prescription drugs, there are numerous examples of patents that have allowed a firm to dominate important markets. One would struggle to find any equivalent example among traditional copyrighted products. Even copyrights on such blockbuster items as Harry Potter or Lord of the Rings, hardly help their holder corner the market for fantasy movies, let alone the broader market for movies aimed at children and young adults. Hence copyright-based market power will usually stem from a significant concentration of copyrighted materials.[52] As such, it might be somewhat easier to detect than patent-based market power, where a detailed understanding of individual patent scope is required.

---

[51] This is only an approximation. A formal analysis would also need to consider the fact that competitive conditions – and thus the profits and deadweight losses associated with a given royalty – are likely to be different in the post-patent period.

[52] Indeed, recent competition law cases involving copyright-based monopoly power have focussed on such concentrations. See for example the AOL-Time Warner merger, the OFT 's review of BskyB or the recent EC review of the acquisition of Telepiu by Stream (i.e. Sky Italy).

Mergers between firms that hold significant intellectual property portfolios should carefully consider the potential effects of an increase in the concentration of ownership of copyrighted material. Such an evaluation is likely to be easier that in the case of patent as the relevant markets in which the intellectual property might confer monopoly power is more readily identified.

The copyright protection of software raises additional issues. Foremost among them is the protection of *interfaces*, i.e. the parts of code that ensure the interoperability between a piece of software and other software packages or peripheral equipment. This as been a crucial aspect of the recent string of competition law cases involving Microsoft. In particular, the issue of interoperability was central to the latest EC complaint where Microsoft was alleged to withhold information necessary to ensure that computer servers using non-Microsoft operating system would work well with personal computers, for which Microsoft 's family of 'Windows' operating systems are dominant.

Reviewing the relevant economic theory on compatibility would require a paper in itself. We will therefore limit ourselves to roughly summarizing its main points. Two situations must be considered depending on whether the products involved are substitute or complements.

The case of rival technologies is discussed in details in Farrell and Katz (1998). The key mechanism is that of *network externalities*. As explained in part I, there are network externalities when the value of a product increases with the number of its users. If two technologies are compatible, then they share the same network: if files can be transferred easily between two word processing packages, then consumers do not care whether more potential co-workers use one of the two. On the other hand, if two rival products are incompatible then the 'network' of each product is limited to its own customers. This creates intense competition aimed at 'building an installed base' of users in order to offer a more attractive product than the rival. If the product or technology purchased are durable goods then consumers will also rely on their *expectation* as to what the network size of the two predicts are likely to be, giving firms huge incentives to manipulate these expectations through product pre-

announcements or, more simply, the diffusion of incorrect information.[53] Incompatibility is also likely to lead to market dominance as rivals might find an early lead in network size to be an insurmountable advantage. Incompatibility also favours incumbents, not only because they might already have a significant installed base but also because consumers might simply expect that a big successful firm entering a new market is more likely to attract a large network of consumers.[54]

Because of this, most economists would agree that overall, ensuring compatibility between rival technologies strengthens market competition. As shown in Farrell and Katz (1998), the effect of compatibility on the firms ' incentives to innovate is more ambiguous. On the one hand, compatibility dilutes the reward from innovation as new consumers attracted by the better product also increase the network size – and therefore the value – of rival products. On the other hand, compatibility opens up R&D competition to firms that could not enter otherwise because of their small expected network sizes.

If technologies are (possibly imperfect) complements, then the openness of interfaces allows consumers to use the two technologies or products together. A natural antitrust concern – strongly evoked in the Microsoft cases - is that a firm with significant monopoly power in one of the two markets would use its control of interfaces to also monopolise the potentially competitive second market. Such deliberate control of interfaces should be seen with great suspicion as a more competitive second market actually enhances the profitability of the firm 's monopoly market. In other words, the benefits from preventing the interoperability of complementary products cannot arise in the original monopoly market. Incentives to prevent interoperability must therefore be found in the firm 's desire to use its position in the monopoly market to also dominate the second market.

Overall, then it seems that opening up interfaces is generally the better policy. One should however remain open to a showing that, in particular cases, ensuring

[53] Disputes between producers of spreadsheet software as to the accuracy of their publishes sales figures is a case in point.
[54] This is not the as the 'reputation' mechanism that underlies the economic theory of trademarks. Here it is sufficient that consumers expect the firm to 'get big' in a new market just because it is already big in another. These expectations do not need to be backed by any objected or expected differences in product quality.

interoperability would have too large a negative effect on innovation incentives. The following principles – adapted from Farrell and Katz (1998) should be helpful in deciding which cases deserve more careful consideration:

• Opening up the interface is likely to be a better policy when the interface itself contributes little relative to obvious ex ante alternatives. In other words, ensuring access to proprietary interfaces is more likely to be innocuous if the interface itself – as opposed to the main body of the software – are not particularly innovative. The other side of the coin is that one should be more suspicious of the motives of the copyright-holder the more arbitrary the interfaces appear to be.[55]

• Forcing the disclosure of sufficient information to ensure interoperability is less attractive when such disclosure cannot be effected without at the same time revealing proprietary information about the innovative aspects of the body of the software itself. However, this argument can also be seen as further support for the claim that such innovative aspects of software products ought to be protected by a patent rather than under copyright law. If they were, then the fact that opening interfaces would also reveal the main innovations would be irrelevant since rivals still could not *use* them without infringing the patent.

c. Trademarks

As discussed in part I, the main economic function of trademarks is to facilitate reputation-building. Like other intellectual property assets, trademarks can be the source of significant monopoly power. What sets trademarks apart is the essential fragility of the advantage that they might confer. While reputations take a long time to build they can be seriously damaged by a single incident. Even the largest firms with the strongest 'brands' are not immune. A few years ago, Coca-Cola took almost a year to get over the fact that a number of European users were apparently affected by a foreign substance found in Coke bottles.[56] Union Carbide never recovered from the Bhopal tragedy and Exxon clearly suffered from the Valdez pollution. This fragility has two main consequences for competition policy.

[55] And the more often they are changed without a substantial technical reason.
[56] See Landes and Posner (1987) for more examples and a similar view.

The first implication is that one should expect firms with strong trademarks to keep a tight control over the production and sales of goods or services bearing their name. If they did not, then the risk would be great that another party allowed to use their trade name would 'free ride' on their reputation to sell shoddy (and therefore cheaper) products at the premium price commanded by the brand 's reputation. Such behaviour could easily ruin the valuable goodwill associated with the brand. Hence competition policies forcing companies to license their trademark or preventing them from including clauses to control the behaviour of their licensee would ultimately be counterproductive as they would remove the firm 's incentive to provide consumers with high quality goods.[57] In our opinion, competition authorities should take a lenient view of refusal to license trademarks as well as of most restrictive clauses found in trademark licensing agreements[58].

The fragility of reputations also affects our view of the practice of *umbrella branding* (also known as *brand extension*) whereby a company uses a trademark made famous by the sale of one product to enter into another market. Perhaps the most famous recent example is *Virgin*, which has now been used to brand products as different as airlines, train services, electricity supply and cola drinks. Other examples abound, from the *Easy*/jet/car name to *Dior*. Such brand extension strategies raise legitimate competition policy issues as a firm is essentially using an advantage acquired in one market to enhance its position in another. We believe however, that there is little room for concern – and even less room for appropriate antitrust intervention. This position is based on two main arguments. The first one is that the 'foreclosure' mechanism described in Whinston (1990) (see above) does not apply. That mechanism relies on the fact that the firm only sells its two products as a bundle. This is not the case with umbrella branding as consumers are still completely free to buy the firm 's original product without also purchasing the new good or service that it offers under the same name. The second argument is that the firm 's brand name is not separable from the quality of the product that it sells *in its new market*. While the

[57] Perry and Groff (1986) also show that trademark licensing reduces the variety of products available to consumers and that, in most cases, this decreases overall welfare.
[58] Franchising agreements often include strict clauses as to how the product should be made and displayed, territorial restrictions aimed at linking products sold back to specific franchisees and 'tie-ins' whereby the franchisee must purchase some inputs from the franchisors.

brand name might provide an initial advantage, it will be useless if the quality of the new product does not meet consumer expectations. In fact, because bad performance would likely also affect the firm 's sales in its original market, the firm will have greater incentives to maintain quality in its new market than if it were not using the same brand name (see Wernerfelt 1988).

These two arguments do not mean that umbrella branding cannot have ant-competitive effects. The very fact that consumers understand that a well-known brand has 'more at stake' will convince them to give it the benefit of the doubt and try its product. This means that, provided the brand-extension firm delivers on quality, a new company offering a product of similar quality might find it hard to get a foothold into the market.[59] Still, it is hard to see what would constitute an effective remedy. Compulsory trademark licensing would run into the difficulties described above: if tight controls are required to preserve quality, can we be sure that the licensee would provide effective competition? One could of course simply prevent a company with a dominant trademark to use it in new markets. However, as we have explained, such a remedy would reduce the firm 's incentive to provide a high-quality product *both* in the new market and in the markets where it is currently dominant.

d. Competition Policy and IPRs in a Small Open Economy

Even though many 'small' economies also tend to be rather open to trade, it is useful to distinguish between the two concepts as 'smallness' and openness have rather different implications for IPRs and their relationship with competition policy.

A. Small

To focus on the effect of 'smallness', imagine that there is no trade in goods whatsoever between countries. However, trade in technologies is allowed. Let us consider a given industry, say cars. An economy can be 'small' in two senses. Firstly, the size of its internal market for cars can be modest compared to the fixed costs that firms must incur to compete in it. Secondly, the internal market can be

[59] On the other hand, a firm already established in that market should have no such problem as it has already demonstrated its quality to customers.

small compared to other markets for cars in the rest of the world. Notice that this two notions of smallness are not necessarily closely related. In an industry where fixed costs are generally low, a country can be small in the second sense without being small in the first sense. Notice as well that both definitions of smallness apply at the industry level. Hence a country can very well be small for one industry and not for the other: Belgium is a small country for cars but not for chocolate.

A country that is small in the first sense will be characterised by highly concentrated product markets. In economic terms, the industries for which the country is small will have a structure of a natural oligopoly – or even monopoly – as only a limited number of firms can possibly survive. However, this does *not* mean that intellectual property is more likely to be a source of market power – and therefore a potential concern for competition policy. Even if the national market were served by a single firm relying on a proprietary technology (its own or obtained through a licensing agreement), there might be several other technologies that could potentially be used to serve the market. The fact that they are not reflects the natural monopoly structure of the downstream market, not a lack of competition in the upstream market for technologies. If anything, the greater downstream concentration found in 'small' markets means that any market power in the upstream market for 'technologies' is more likely to be somewhat limited by countervailing 'buyer power' downstream.

Smallness can only increase the market power of IP holders if there is a country-specific (fixed) cost to using the intellectual property. One such cost is the cost of obtaining legal protection. This suggests that 'small' countries should take special care in not making their intellectual property protection process too onerous. Country-specific costs also arise if the technology needs to be adapted in order to meet local needs. So small countries with local conditions that differ significantly from the average are more likely to experience high concentrations in the 'upstream' IP market.[60]

[60] In some cases, technologies might not be introduced because these adaptation costs are too high compared to the expected private benefits from serving the small market even though their introduction would increase welfare. In such situations, subsidies might be called for. As trade agreements often frown on industry-specific subsidies, small countries might want to ensure that the appropriate escape clauses are included.

Being small in our second sense affects the effectiveness of the country 's policies, including competition policy and intellectual property policy. As we have discussed before, intellectual property law strikes a balance between the need to provide inventors with adequate rewards and the desire to promote the widest possible use of the invention and to ensure quick diffusion of the embodied knowledge. As the markets for many inventions tend to be global, the incentives given to researchers depend on the total returns – and thus on the type of protection – that they can enjoy in a large number of countries. If a country is small, then the effect of its own IPR protection is negligible. On the other hand, the country still fully enjoys the benefits of being able to use the invention and/or consume the resulting products. This means that, in a global economy, the traditional IP Law trade-off is, from a single country 's perspective, biased against the 'reward' factor. In other words, individual countries have incentives to offer low level of IPR protection, effectively free-riding on the protection offered by others. This temptation to free-ride would be especially strong for a small country since its own policy has essentially no effect on worldwide incentives to invent. Two factors can mitigate the small country 's tendency to under-protect intellectual property. Firstly, as mentioned above, the country might not be able to enjoy the fruits of technology unless it is adapted to the local conditions. Getting firms to undertake such adaptation clearly requires some level of local IPR protection.[61] Secondly a 'small' country might nevertheless be relatively 'big' in terms of its own inventive activity. The larger the country 's share of 'inventors' the greater its incentive to provide significant IPR protection.

The previous argument explains why smaller (and less innovative) countries might be more reluctant to offer strong IPRs. However, one should not be too quick to give it a normative interpretation. As we will see below, there are good reasons for small countries to still have strong IP laws, once international trade considerations are taken into account.

B. Open

Lets us now assumed that the country is completely open to trade. The most obvious consequence of this openness is that relevant markets are now more likely to

[61] See Diwan and Rodrik (1990) for a formalisation of this argument.

extend beyond the national boundaries, making it less likely that there will be significant monopoly power in product markets.[62] On the other hand, allowing for free trade in goods does not have any direct effect on the expected level of concentration in IP markets. The only implication for the relationship between antitrust and IP Law is therefore that IPR holders with market power are somewhat less likely to face significant buyer power in the relevant downstream markets.

Openness has more important consequences for the design and implementation of IPR Laws themselves. A first consideration is that open economies make it harder to enforce national IPR Laws: even if a country rigorously enforces IPRs at home, it can do little to ensure that effective protection is also given in foreign markets. This is especially true of small countries as a large economy can hurt infringers by excluding them from a valuable market while a small one cannot. Hence small *innovative* economies might find it profitable to enter and respect international agreements on IP protection. A second factor is that trade makes innovative nations of all sizes more concerned about infringement. With closed economies, any loss of know-how in a foreign country only has adverse consequences in that market. With free trade, the damage can easily spread to other markets. This explains why large, innovative countries ask for guarantees on IP protection when entering into trade agreements with smaller economies, especially if these are perceived as less innovative[63].

Small open economies depend very much on international trade for their economic success. This provides two further reasons why, despite the free- riding argument outlined in section A, small countries might still want to have strong IPR laws and enforce them strictly. Firstly, they are very vulnerable to retaliation from large trading partners, both on the IPR and trade fronts. Secondly, if a small economy wants to further develop its innovative sector and strengthen its companies on the international scene, it might as well get them accustomed immediately to the kind of strict(er) IP Law enforcement that they will face in some of the world 's more lucrative markets.

[62] The 'experiment' we are considering is one where, for given levels of fixed costs, the size of the market increases.
[63] For more discussion of these and other related issues see Sotchmer (2002).

C. Conclusion

The main conclusion of this section is that being a small open economy does not have major implications for the relationship between IPR Law and competition policy. This means that the general arguments presented in this paper still apply. As for the design of IP Law itself, small countries have – in theory – a strong incentive to offer lower levels of protection than larger economies. In practice though, a small open economy – especially one with significant expertise in innovation, might still find it best to adopt IPR policies that are similar to those of its large trading partners.

4. Conclusion

We have tried to provide a consistent economic view of the legal protection of intellectual property and its interface with competition policy. We made two main points. Firstly, the protection granted under patent, copyright and trademark laws is broadly consistent with the principles of the economic theory of property rights. In particular, the specificity of intellectual property assets (as opposed to other forms of investment) is accounted for in the type of property rights attached to them. This leads to our second conclusion: since the main distinguishing features of various types of assets are handled effectively by their respective property rights regime, all sources of monopoly power should be treated equally under Competition Law. In particular, international property should not receive special consideration because of its 'contribution to the creation and diffusion of knowledge'.

Equal treatment of monopoly power, regardless of its source does not mean that the interface between intellectual property and competition law does not have any special characteristics. In fact, we argue that some types of (potentially) abusive practices are more likely to arise in the presence of IP-based dominance. We then provide an economic analysis for a number of these practices, including cross-licensing, patent pools, grant-backs the copyright protection of software interfaces and

the practice of ‘umbrella’ branding. We also address the role of intellectual property portfolios in merger reviews.

References

Anderman, S.D., 1998, *EC Competition Law and Intellectual Property Rights*, Oxford: Oxford University Press, 320 p.

Ayres, I. and P. Klemperer, 1999 "Limiting Patentees ' market Power Without Reducing Innovation incentives: The Perverse Benefits of Uncertainty and Non-injuctive Remedies", *Michigan Law Review*, Vol. 97, pp. 985 – 1033.

Bakos, Y., E. Brynjolfsson and D. Lichtman, 1999, "Shared Information Goods", *Journal of Law and Economics*, XLII, pp.117 – 155.

Bernheim, D. and M.D. Whinston, 1990, "Multimarket Contact and collusive Behaviour", *Rand Journal of Economics*, Vol. 21, pp. 1 – 26.

Choi, J.P., 2002, "A Dynamic Analysis of Licensing: The "Boomerang" Effect and Grant-back Clauses", *International Economic Review*, 43:3, pp. 803 – 829.

Cornelli, F. and M. Schankerman, 1999, "Patent Renewals and R&D Incentives", *Rand Journal of Economics,* 30:2, pp. 197-213.

Dealessi, L., and R. Staaf, 1994, "What Does Reputation Really Assure – The Relationship of Trademarks to Expectations and Legal Remedies", *Economic Inquiry,* 32:3, pp. 477-485.

DeNicolo, V., 2000, "Two Stage Patent Races and Patent Policy", *Rand Journal of Economics,* 31:3, pp. 488-501.

Diwan, I. And D. Rodrik, 1990, "Patents, Appropriate Technology and North-South Trade", *Journal of International Economics*, 29 (3-4), pp. 293 – 310.

Erdem, T., 1998, "An Empirical Analysis of Umbrella Branding", *Journal of Marketing Research,* 35:3, pp. 339-351.

Erdem, T. and V. Sun, 2002, "An Empirical Investigation of the Spillover Effects of Advertising and Sales Promotions in Umbrella Branding," *Journal of Marketing Research*, 39:4, pp. 408-420.

Eswaran, M., 1994, "Cross-Licensing of Competing Patents as a Facilitating Device", *Canadian Journal of Economics*, 27:3, pp. 698 – 708.

Farrell, J. and M.L. Katz, 1998, "The Effect of Antitrust and Intellectual Property Law on Compatibility and Innovation", *working paper*, University of California – Berkeley, Department of Economics, April 1998.

Fershtman, C. and M.I. Kamien, 1992, "Cross-Licensing of Complementary Technologies", *International Journal of Industrial Organization*, 10,

pp. 329 – 348.

Gallini, N.T., 1992, "Patent policy and Costly Imitation", *Rand Journal of Economics*, 23, 1, pp. 52-53.

Eswaran, M., and N. Gallini, 1996, "Patent Policy and the Direction of Technological Change", *Rand Journal of Economics,* 27:4, pp. 722-746.

Gilbert, R.J. and Shapiro, C., 1990, "Optimal patent Length and Breadth", *Rand Journal of Economics*, 21,1, pp.106-112.

Green, J.R. and S. Scotchmer, 1995, "On the Division of Profit in Sequential Innovation", *Rand Journal of Economics*, 26,1, pp. 20-33.

Harhoff, D., P. Régibeau and K. Rockett, 2001, "Some Simple Economics of GM Food", *Economic Policy*, 33, pp. 263-299.

Horstman, I., G. MacDonald and A. Slivinski, 1985, "Patents as Information Transfer Mechanisms" *Journal of Political Economy*, 93, pp. 837-858.

Hui, K-L. and I.P.L. Png, 2002, "On the Supply of Creative Work: Evidence from the Movies", *American Economic Review,* 92:2, may, pp. 217-220.

Kamien, M.I. and N.L. Schwartz, 1982, *Market Structure and Innovation*, Cambridge University Press, 241 p.

Katz, M. and C. Shapiro, 1985, "On the Licensing of Innovations", *Rand Journal of Economics*, Vol. 16, pp. 505-520.

Klemperer, P., 1990, "How Broad Should the Scope of Patent Protection Be?", *Rand Journal of Economics*, 21,1, pp. 113-130.

Kremer, M., 1998, "Patent Buyouts: A Mechanism for Encouraging Innovation", Q*uarterly Journal of Economics,* CXIII:4, pp. 1137-1167.

Landes, W.M., 1992, "Copyright Protection of Letters, Diaries, and Other Unpublished Works: An Economic Approach", *Journal of Legal Studies*, XXI, pp. 79 – 113.

Landes, W.M. and Posner, R., 1987, "Trademark Law – An Economic Perspective", *Journal of Law and Economics,* 30:2, pp. 265-309.

Lane, W., 1988, "Compulsory Trademark Licensing", *Southern Economic Journal,* 54:3, pp. 643-655.

Lanjouw, J.O. and M. Schankerman, 2002, "Enforcing Intellectual Property Rights: Suits, settlements and the Explosion in Patent Litigation", *mimeo*, Department of Economics, Yale University.

Manelli, A., P. Régibeau and K. Rockett, 1994, "Trademark Protection and Post-

Patent Competition", *Cuadernos Economicos*, 57:2, pp. 115-139.

Matutes, C., P. Régibeau and K. Rockett, 1996, "Optimal patent design and the Diffusion of Innovation", *Rand Journal of Economics*, 27,1, pp. 60-83.

Nordhaus, W., 1969, *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*, Cambridge, MA: MIT Press.

Novos, I.E. and M. Waldman, 1984, "The Effects of Increased Copyright Protection: An Analytic Approach", *Journal of Political Economy*, 92, pp. 236 – 246.

O' Donoghue, T., 1998, "A Patentability Requirement for Sequential Innovation", *Rand Journal of Economics*, 29,4, pp. 654-679.

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

Perry, M.K. and Groff, R., 1986, "Trademark Licensing in a Monopolistically Competitive Industry", *Rand Journal of Economics*, 17:2, pp. 189-200.

Png, I., 1995, "Why are Some Products Branded and Others Not", *Journal of Law and Economics,* 38:1, pp. 207-224.

Régibeau, P., 2000, "La Protection des Inventions", in P. Pestieau and B. Jurion (Eds), Finances Publiques, Finances Privées, Les Editions de l' Université de Liège, pp.143 – 167.

Schechter, R.E., 1986, *Unfair Trade Practices and Intellectual Property*, West Publishing Co., St Paul, Minnesota, 272 p.

Scherer, F., 1972, "Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation", *American Economic Review,* 62, pp. 422-427.

Scotchmer, S., 1999, "On the Optimality of the Patent Renewal System", *Rand Journal of Economics,* 30:2, pp. 181-196.

Scotchmer, S., 2002, "The Political Economy of Intellectual Property Treaties", *Journal of Law, Economics and Organisation,* forthcoming.

Scotchmer, S., and J. Green, 1990, "Novelty and Disclosure in Patent Law", *Rand Journal of Economics*, 21, pp. 131-147.

Scotchmer, S., and S. Maurer, "The Independent Invention Defense in Intellectual Property", *Economica,* 69, pp. 535-547.

Scotchmer, S., and S. Maurer, 2004, "Profit Neutrality in Licensing: The Boundary Between Antitrust Law and Patent Law", *Competition Policy Center, University of California, Berkeley, Working Paper CP04-43.*

Shapiro, C., 2003, "Antitrust Limits to Patent Settlements", *Rand Journal of Economics*, Vol. 34, 2, pp. 391 - 411.

Takeyama, L.N., 1997, "The Intertemporal Consequences of Unauthorized Reproduction of Intellectual Property", *Journal of Law and Economics*, Vol. XL, pp. 511 – 522.

Takeyama, L.N., 1994, "The Welfare Implications of Unauthorized Reproduction of Intellectual Property in the Presence of Demand Network Externalities, *Journal of Industrial Economics*, Vol. 58, pp. 155 – 166.

Tandon, P., 1982, "Optimal Patents with Compulsory Licensing", *Journal of Political Economy,* 90:3, pp. 470-486.

Van Dijk, T., 2000, "License Contracts, Future Exchange Clauses and Technological Competition", *European Economic Review*, 44:8, pp. 1431 – 1448.

Veall, M., 1992, "Brand/Product Innovation and the Optimal Length of Trademark/Patent Protection", *Economics Letters,* 40:4, pp. 491-496.

Wernerfelt, B., 1988, "Umbrella Branding as a Signal of New Product Quality – An Example of Signalling by Posting a Bond", *Rand Journal of Economics,* 19:3, pp. 458-466.

Whinston, M.D. (1990), 'Tying, Foreclosure, and Exclusion', *American Economic Review*, 80:4, September, pp. 837 - 859.

Wright, B., 1983, "The Economics of Invention Incentives: Patents, Prizes and Research Contracts", *American Economic Review,* 73:4, pp. 691-707.