CHAPTER 2

*The Law and Economics of International Intellectual Property: A Primer**

Jonathan Putnam

Institute for Policy Analysis, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada
CRA International, 80 Bloor Street West, Suite 1501, Toronto, ON M5S 2V1, Canada
E-mail address: jon.putnam@utoronto.ca

Abstract

I begin with a dispute over a fox hunt, by which to understand the law of tangible property, then develop that metaphor for the major types of intellectual property. I start with domestic U.S. patent law for the sake of concreteness, and generalize to other jurisdictions and types of intellectual property. In the latter parts of the paper I discuss the international implications of intellectual property, including especially the effects of information spillovers. The last part of the paper describes the hazards in analogizing "trade" in intellectual property rights to trade in goods, and particularly in interpreting international patent data. These hazards motivate the search for a structural model specially adapted to the purpose of valuing international intellectual property rights and rules. The goal is to give economists a simple and integrated framework for analyzing intellectual property across time, jurisdiction and regime type, with an eye towards eventually developing other incentive systems that have the advantages of property (such as decentralized decision-making), but fewer of the disadvantages.

Keywords: Patents, intellectual property, international trade

JEL classifications: F02, K11, K33, O34

---

* Many of the ideas in this paper were developed under the auspices of the ORDCF Professorship in the Law and Economics of Intellectual Property, Centre for Innovation Law and Policy, University of Toronto. I thank the students and colleagues who taught me about the law of property, particularly Dean Ron Daniels, at whose insistence I learned the law of property in Canada. Parts of this paper are abridged from "International Patenting," which develops empirical methods of analyzing the international patent system. The mistakes are mine.

FRONTIERS OF ECONOMICS AND GLOBALIZATION
VOLUME 2 ISSN: 1574-8715
DOI: 10.1016/S1574-8715(07)00002-4

1. Introduction

Given its various dimensions (law/economics; national/international regimes; the four main types of intellectual property), the paper's title implies 16 analytical combinations. That alone makes any comprehensive treatment of the issues infeasible. Instead, I reason by analogy – from tangible to intangible property law, from patents to other types of intellectual property instruments, and from international goods and services trade to international trade in intellectual property rights. My objective is to help economists better appreciate intellectual property law, thereby to create better economic models of behavior and legally implementable prescriptions for policy reforms.

1.1. Law and economics

The "law and economics" of international intellectual property rights is something of a misnomer. On the continuum from micro- to macro-economics, the "micro" end is anchored to some kind of optimizing behavior (in a single-agent economy) or exchange (in a multiple-agent economy). Without optimization of some kind, economics has little to say. If the elements of a classical microeconomy are "atomistic" production and exchange, under conditions of convexity and continuity that lead to equilibrium, then property rules play the part of "sub-atomic" forces, which not only shape individual behavior and transactions, but also select the sample of transactions that underlie the equilibrium. In other words, on the micro–macro economic continuum, property law is located at a *sub*-micro point: property rules may facilitate optimization and exchange, but they are social "primitives" whose function is legally independent of, and logically prior to, optimizing behavior.

At least when viewed in this way, there is little true intersection between the literatures of economics and of international intellectual property law. For that reason, I make no attempt to review the "law and economics literature." But each literature has natural entry points for interested outsiders. Economists who desire a more complete understanding of property law should consult a leading economics-oriented legal textbook, such as Dwyer and Menell (2001). For a law and economics-based introduction to U.S. patent law, see Merges and Duffy (2002). Chisum (serial) is an exhaustive and up-to-date treatise of the legal (U.S.) frontier. Analogous texts and treatises include: Merges *et al.* (2006) (general intellectual property); Nimmer (serial) (copyright); Gilson and LaLonde (serial) (trademark), and Milgrim (serial) (trade secrets). A useful guide to online research is Weigmann (2000).

For industrial organization-oriented economists, Hovenkamp *et al.* (2002) reviews the history of applying the rules of antitrust law to intellectual property. The international legal dimensions of intellectual property, including the role played by international agreements and institutions, are covered comprehensively in Goldstein (2001) and Dinwoodie *et al.* (2001, 2003). A standard reference for international trade regulation is Trebilcock and Howse (2005); their Chapter 13 analyzes the effect of trade regulation on intellectual property rights (see also Jackson, 2002). Lerner (2002,

2005) provide a useful review of the economically relevant evolution of patent laws and practices, respectively, across a large number of countries over a long period.

Going the other way, lawyers interested in economic analysis using intellectual property rights will gain an extensive overview from the papers in this volume, and from related papers by the same authors, who have produced a large share of the leading economics scholarship. An especially good review of intellectual property rights as incentive mechanisms in a single-country context is Gallini and Scotchmer (2002). Like many other economists, Gallini and Scotchmer discuss the economic interpretation of "broad" vs. "narrow" patent protection, and the interaction between patent breadth and licensing, at a level of abstraction which economists find useful, but which lawyers will immediately recognize as insensitive to national rules and to differences in national regimes. Similarly, the multi-dimensional concept of legal "imitation" becomes a one-dimensional parameter in leading international trade-based papers such as Eaton and Kortum (1996, 1999). Despite the force of their economic ideas, these papers lack much legal nuance.[1]

In general, economists count intellectual property rights, as "economic indicators" (Griliches, 1990), generally without much thought for the multiple ways in which they alter agents' opportunity sets, and almost always without regard for their prices (but see Lanjouw *et al.*, 1998). As Section 6.2 shows, even relying on simple quantities often results in ambiguity, and in the misinterpretation of international intellectual property data. In such cases, an appreciation of the legal details greatly facilitates the creation of informative structural models.

Given the opposite ends of the continuum from which they approach the analytical issues, international lawyers will find the legal analysis truncated and parochial, while macro/international economists will find it idiosyncratic and arcane. The paper will have done its job if it persuades economists to tailor their models to the peculiarities of intellectual property acquisition and "trade," and if it persuades lawyers that such models can yield reliable economic insights into the effects of their idiosyncratic rules and institutions.

1.2. National and international law

Until recently, international intellectual property law functioned primarily to impose some principles of standardization and non-discrimination upon national laws. Lack-

[1] To take but one example: several studies (e.g., Deardorff, 1992; McCalman, 2005) conclude that, in a stylized model of trade between developed (North) and undeveloped (South) regions, the North benefits from "stronger" global intellectual property rights, while the South either does not benefit, or is harmed. To economists, this observation helps account for the "strengthening" of intellectual property rights in the Uruguay round of the General Agreement on Tariffs and Trade (GATT). But, apart from this basic welfare conclusion – the value of which should not be underestimated – there is little or no discussion of what "stronger" intellectual property rights means, either as a matter of the domestic law that must create and enforce property rights, or the internationally standardized law to which GATT-member countries bind themselves.

ing any independent enforcement mechanism, international law was primarily a collection of national laws, stitched together. For that reason, any study of international law depended primarily on understanding the constituent national regimes.

That is still mostly true. But with the advent of the Trade-Related Aspects of Intellectual Property Agreement ("TRIPS"), international intellectual property law has become more fully integrated into international trade regulation. As a legal matter, the failure to comply with international IP standards is punishable through the international trade enforcement mechanisms inherited from GATT and amplified in the World Trade Organization (WTO). Since there has been a secular increase in the nominal standardization of national IP laws, even before TRIPS, the primary economic effect seems to be that states exhibiting lax actual enforcement of their nominal standards now face the prospect of real, concerted trade retaliation.

The main critique of the new regime, as elucidated in the North–South models described in footnote 1, is that it does not allow member states to tailor their IP laws to their individual economic circumstances, with potentially deleterious effects on welfare. To appreciate this debate, one must understand the various dimensions of national law, including the dimensions in which standardization allegedly has been harmful. For this reason, I focus on national IP laws as the constituent elements of the international regime. I illustrate most of the ensuing discussion with reference to U.S. property rules, particularly U.S. patent rules, not because those rules are legally or economically superior, but because they are familiar and (at the level of abstraction presented here) widely shared. They function as a baseline, not a norm.[2]

1.3. Types of intellectual property

Different intellectual property regimes cover different informational subject matter, using different rules. In the law, these rules tend to evolve independently; there is very little common legal or economic theory underlying the several types, beyond such basic platitudes as "unrestricted access to others' information leads to underinvestment in new information." At the risk of some oversimplification, I have tried to take a unified approach to the economic attributes of each IP system. In particular, Table 1 summarizes some of these attributes. In addition to comparing regimes, this approach also illustrates the parameter space with which to compare a single regime across countries, and to evaluate some of the effects of international standardization.

---

[2] The reader should note that, in addition to differences among common law jurisdictions, of which the United States (as a former British colony) is one, these jurisdictions differ from civil law jurisdictions, of which continental western Europe and its colonial progeny are the most important examples. Still other countries, such as Japan and members of the former Communist bloc, have grafted Western-style intellectual property laws onto their indigenous tangible property laws, so the relationship between the two is different than is described here. While the idiosyncratic differences among the national laws of tangible property are often critical, and may have important economic consequences, for individual patentees, they lie beyond the scope of a primer on the law and economics of international intellectual property.

Table 1. Comparison of economically relevant features of intellectual property systems, by type

Feature	Patent	Copyright	Trademark	Trade secret
Subject matter	Apparatus, process, composition	Literary and audio/visual works	Identifying mark	"Anything. . ." / commercial processes
Broad product attribute that receives protection	Function	Form	Sign	Scarce information
Title	By examination	By creation	By examination (federal); by distinctive use in commerce	By creation
Possession	Reduction to practice	Expression in a tangible medium	Distinctiveness	Ongoing reasonable efforts to maintain secrecy
Priority	Absolute novelty	Originality	Distinctiveness	None
Term	20 years	Life +50–70 years	Indefinite, subject to use	Indefinite, subject to secrecy
Owner has the right to use	No	No	Yes	Yes
Notice at creation	Claims, as interpreted by specification	None	Publication of mark	None; may be specified contractually
Notice at enforcement	Marking	Marking; registration	Marking	None
Increment to public information stock	Enabling disclosure	Idea	Non-commercial significations	None
Independent creation is a defense to infringement	No	Yes	No	Yes
Injunctive relief	Yes	Yes	Yes	Yes, subject to ongoing secrecy
Public use exceptions	Limited research; compulsory license	Fair use; compulsory license	Non-commercial, non-disparaging use	None
Public domain exclusions	Discoveries	Facts	Non-commercial references	Reverse engineering
Antitrust risk	High	Medium	Low	Low to none
Criminal liability	No	Yes	Yes	Yes
Intent determines liability	No (may affect damages)	No (civil); yes (criminal)	Yes?	Yes

If the 16 potential dimensions of the paper were consolidated into one theme, it would be this: just as the non-rivalrous nature of information complicates the international analysis of the economic good, so the non-rivalrous nature of intellectual property complicates the international analysis of the legal rights. I emphasize the importance of that complexity, the better to manage and improve upon it.

The purpose of this paper, then, is threefold: to explain how sub-microeconomic property forces arise generally in national law (Section 2); to describe their role in intellectual property, focusing in particular on their economic effects (Sections 3 (patents) and 4 (other intellectual property)); and to review some of the ways that introducing an international dimension complicates intellectual property, both for lawyers (Section 5) and for international trade-oriented economists (Section 6).

If nothing else, this paper illustrates the richness of the legal treatment of information. Although intellectual property laws exist in part to solve the classic problem of underinvestment in information (a public good), intellectual property rarely implies the true ownership of information. Thus, for economists, it is critical to understand what "ownership" means, and how the law views "information." For that, one must understand something about tangible property.

2. *Property*

Traditional microeconomics is primarily concerned with the production, consumption and exchange of discrete quantities of goods. Economists generally take for granted the ownership rights necessary to produce, consume or exchange each unit of a good, or to dispose of it. But property law is concerned with the constituent rights of ownership, such as the right to use a unit, or the right to exclude others from using it, or the right to sublicense one of these rights to others, whether or not the right holder is the legal owner. In general, economics treats "ownership" as being a primitive, homogeneous state, while property law treats "ownership" as the sum of a set of constituent rights. When rights less than full ownership do arise in economic analysis, their enumeration and allocation are not generally the focus.[3]

2.1. "Rights, not things"

"Property" is the set of rules by which society adjudicates competing claims to a resource. These rules may facilitate private market activity, such as resource transfers, and they may promote socially valuable investments, such as coordinated land development. Some property rules are also efficient, at least in the static sense that

[3] For example, in the analysis of agricultural share tenancies, analysts explain the efficiency properties of the rental contract between landlord and tenant (which grants the tenant certain rights less than ownership) in principal–agent terms, or in terms of sharing extrinsic risk. They are much less likely to ask whether and why, for example, a land lease conveys to a tenant the right to exclude a landlord, despite the landlord generally retaining all the rights of an owner, or whether the tenant has a sub-lease right.

they produce a Pareto-optimal allocation. But similar property rules are found in non-market as well as market economies; they are not, in themselves, market-based.

Market-based activity derives from voluntary exchange. In law, voluntary exchange is accomplished most widely through contracts, most often between two parties. By enforcing private contracts, the government binds agents to their voluntary promises. The resulting market allocation represents the fixed point of these bilateral exchanges.

As a social, or political, compact, property governs a property holder's economic relations, not with those with whom he has contracted privately, but with "all others." Property laws can therefore be thought of as a set of social default rules for determining resource access and control. By enforcing property laws, the government binds agents to a web of implicit mutual promises, when bilateral contracts themselves are absent or incomplete. Almost surely, the resulting allocation is not an equilibrium. Rather, it helps define an endowment from which the movement to equilibrium can proceed.

2.2. The property rights bundle

One of the first lessons property lawyers learn is that "property is rights, not things." Although non-lawyers routinely confuse the two (as in, "Get off my property"), the legal distinction is essential, but not always obvious.

"Property" is a collection of rights. In neoclassical economics, "property rights" generally means "ownership" – the fullest set of rights available. For example, I can exchange guns for butter because I own the guns. I rely on my ownership interest in the guns, which includes the rights to transfer each of them individually, when transferring my interest to the owner of the butter. Traditionally, the only interesting economic issue is the rate of exchange.

If, however, I lacked the right to transfer guns, and had to acquire that right from someone else (say, the owner of the butter), then the rate of exchange of the *goods* becomes inextricably linked to the exchange of *rights*. Similarly, the existence of competitive equilibrium depends on the axiom of free disposal – each agent's capacity to rid himself of units that do not generate positive marginal utility. But disposal can only be free if the owner also possesses the *right* to dispose. Though economists take these rights for granted, lawyers know that the right to transfer, and the right to destroy or dispose, are just two of the many distinct rights that may be bundled in "ownership."

The most essential of property rights is the right to exclude others. This means, for example, that if you erect a building on one inch of my land, I can insist that you remove it, and the government commits to me that it will force you to remove it. In general, nothing in the law requires a cost-benefit calculation of the exercise of this right. In particular, nothing requires me to accept compensation, or otherwise to contract with you, in lieu of actual removal.[4]

---

[4] From the point of view of remedies, this is the most fundamental difference between contract and property law. The "remedy at law" for infringement of a contractual right is monetary compensation equal to

In addition to the rights to exclude, transfer and destroy, there are many others: the right to use, the right to the yield (of land), the right to minerals (under land), the right to air or light (over land), the right to sublicense rights to others, etc. The set of available rights depends in general on the ways in which one may extract benefit from the resource, but rights exist independently of whether they have economic value.

In general, property rights apply against "all others," but this class may be narrowed by contract or by the operation of other property rights. So for example, I may exclude all others from my land, but the law will imply, from my contracting for municipally metered water, the existence of a license to the water company to enter my land to read the company's meter. This purely fictional "bargain" grants to the water company the right to use my land for a specified time and purpose. When the right to use my land does not depend on a contract (in this case, the contract for water), but survives the termination of the contract and my transfer of the land to someone else, then it is a property right, which is said to "burden" the land. An easement is one such burden. Thus, property rights to the same resource – in particular, the right to use – may be distributed among multiple users, even though in general there can be only one owner.

One might suppose that when the resource in question is land, it is fairly simple to define the "thing" to which property rights attach.[5] As any property law student will attest, disputes over what "land" means and what rights attach to it have assumed almost infinite variety over the ages, so this supposition is misguided. But it turns out that defining the "thing" protected by intellectual property rights is still more subtle and varied. In particular, despite the connotations of "intellectual" property, the law generally does not protect information *per se*.

One way of mitigating some of the disputes over resources is to create a system of explicit "titles." Title defines the resource in question using unambiguous terms, such as by reference to a publicly agreed scheme for dividing large plots of land into smaller units having defined boundaries. Title also associates each unit with a publicly identified titleholder.[6] The public nature of a title system serves an important social objective of property law: it provides notice to all others – including potential competing claimants and potential purchasers – of the titleholder's identity, and of the current definition and status of the titleholder's claims (including impairments and encumbrances, such as a lien).

Two other general aspects of property law are relevant to the ensuing discussion. First, in the absence of explicit title, "title" is a relative concept. The traditional job of

---

the loss, but the "equitable remedy" for infringement of a property right is to restore the property owner's right to exclude by actually excluding the infringer from future infringement. The government accomplishes this remedy by issuing an injunction, which obligates the infringer to certify to the government that it has complied with the exclusion.

[5] Property rights in land are "real property," not because land is tangible but because, under English law, realty derived from royalty, i.e., from the *realm* of the King. "Chattel" is property in tangible resources not permanently affixed to land (from "cattle").

[6] The English title system – the Duke of York, etc. – granted defined tracts of land to titleholders, in exchange for political and military service to the King. One of the rights granted to titleholders was the subsidiary right to grant under-titles to smaller tracts, in exchange for similar service ("Earl of Kent").

the common law is to determine which of two claimants has better title, not who of all possible claimants has the best title. Thus, if I find a wallet and you take it from me without permission, I have better title than you, and as between us the law will take the wallet from you and give it to me. But it is not "my" wallet, in the sense that the original owner may, at some point in the future, deprive me of the wallet, even if the original owner was careless and even if I devoted considerable resources to finding the wallet. As the finder, my right to exclude "all others" means "all others with an inferior claim."[7]

Second, property rules create *ex ante* investment incentives, to the extent that they apply prospectively, to potential claims. And to the extent that the resolution of future disputes confirms present investment decisions, property rules support rational expectations. But because property rules implement static, non-price allocations, and they do so retrospectively, they make no pretense of establishing an optimal dynamic investment program, even if they are statically Pareto-efficient. As with any other investment, an investor who externalizes the benefits or costs of the investment lacks the correct incentives to invest.[8] But "incorrect investment incentives" is an economic abstraction; made without reference to an unspecified social welfare function, it generally has no legal significance. Thus, river pollution may be inefficient, but it is not illegal unless someone has an enforceable property interest in clean water. Similarly, the use of information produced by another (copying an author's words, or copying a restaurant chain's choice of locations) may be inefficient, but it is not a misappropriation unless the other has an enforceable property interest in that use.

This brief review underscores the point made at the outset: the legal operation of property rules proceeds without prices, although property rights can be unbundled and priced separately. While efficiency may be a by-product of property rules, it is not a necessary objective or constraint.

2.3. Legal procedure and economic methodology

A large part of the evolution of the law, especially the law of property and especially in common law countries, derives from the resolution of specific disputes over particular facts. This inductive method economizes on judicial resources, and preserves flexibility in the face of technical and institutional change, by obviating the need to formulate

---

[7] In addition, a finder and prospective keeper assumes a duty to attempt to notify the original owner. Thus, finding is not costless, and imposes on the finder the inherently conflicting roles of agent (of the original owner, to reunite him with the wallet) and principal (for herself and her own contingent interest in the wallet, should she be unable to find the original owner). The law typically resolves this conflict by mandating "reasonable" efforts, and by deeming the wallet abandoned (thereby voiding the potentially superior claim) after a "reasonable" (perhaps statutorily defined) period of time.

[8] As in the Coase Theorem, these incentives can be "corrected" statically, to the extent that property rights enforcement requires internalization of the externalities, and as in the Coase Theorem, there may be multiple, economically equivalent means of accomplishing that internalization.

a general theory of the law. The inductive method also leaves room for uncertainty about how to resolve different disputes arising under different facts. As disputes arise and courts decide them, the body of precedent increases, and interstices in the law diminish.

In economic terms, the history of property law is path-dependent, to the extent that it depends on the sequence in which disputes arise. Certain property rules and rights may be path-independent, at least within a given constitutional framework, and these might be termed the "theory" of property law. But, as will become apparent, there is no bright line between the "timeless truths" of property law and the precedent of a particular time and place.

For various reasons, legislatures intervene in this evolution, by passing a statute. They may reverse the conclusions of courts that appear to have erred, anticipate situations that prior disputes have not resolved, or solve coordination failures that cannot be addressed in a dispute between two private parties. Statutes codify, correct, restate, and advance the common law. Much of the law related to intellectual property is statutory law. But statutory law is nevertheless interpreted and applied using many of the legal tools of the uncodified common law. For this reason, it is useful to base an economic interpretation of intellectual property law on an appreciation of the common law of property, and the mechanisms of its evolution.

Because economics proceeds deductively, from first principles, while property law arises inductively, from specific disputes, the two disciplines are apt to collide as they travel towards each other from opposite ends of the general–particular continuum. The interface of that collision is the "law and economics" of property. Economists, especially those that disdain anecdotes, must bear this collision in mind when articulating a "theory of property."[9]

---

[9] For example, a large part of the intellectual force of the general proposition known as the Coase Theorem derived from setting it in the context of a specific dispute: when a steam engine is apt to set fire to a corn field as it passes through, and the farmer threatens to withdraw permission to pass, does it matter to the equilibrium outcome whether the railroad or the farmer has a property right against the other? This example spawned several generations of scholarship on the static allocative effects of the initial endowment of rights. The theorem achieved generality partly because of the distinctiveness of the illustration.

Yet by its very nature that same illustration begs many questions that are critical, not only to the evolution of intellectual property law, but to the primary objective of that law: the stimulation of otherwise underfunded investment. In the absence of considerably more structure, the Coase Theorem says little about whether the initial endowment of property rights affects: the railroad's choice of route; the relative incentives for railroads (farmers) to invest in new spark-arresting (fire-retarding) technology; or the social optimality of the aggregate level of investment.

Whereas economists could take comfort in the claim that the allocation of property rights does not determine the static equilibrium – therefore they could remain indifferent to how the law allocated these rights – it is unlikely that the same can be said with respect to the dynamic equilibrium. If the "anecdotes" of intellectual property law influence dynamic economic equilibrium, then economists must understand the mechanisms of that influence in order to predict the equilibrium and to assist the law in attaining its economic objective.

2.4. Pierson v. Post

To better understand the statutory framework of intellectual property law, and its origins in the common law of tangible property, it is useful to begin with a particular anecdote, in the form of an early American case, *Pierson v. Post* (1805). On its surface, the case demonstrates the manner in which an unpossessed resource comes to be regarded as property under the common law, as well as the interaction between the creation of rights and their exercise. Beneath that surface, *Pierson v. Post* articulates many of the principles that underlie both intellectual property disputes and intellectual property policy. The case also illustrates the limitations of private property law when employed to accomplish explicitly social objectives, like "progress."

The facts, and the legal question they present, are as follows:

> Post, being in possession of certain dogs and hounds under his command, did, "upon a certain wild and uninhabited, unpossessed and waste land, called the beach, find and start one of those noxious beasts called a fox," and whilst there hunting, chasing and pursuing the same with his dogs and hounds, and when in view thereof, Pierson, well knowing the fox was so hunted and pursued, did, in the sight of Post, to prevent his catching the same, kill and carry it off . . .
> The question submitted by the counsel in this cause for our determination is, whether Lodowick Post, by the pursuit with his hounds in the manner alleged in his declaration, acquired such a right to, or property in, the fox as will sustain an action against Pierson for killing and taking him away?[10]

Although the facts may seem idiosyncratic to a particular time, place and production technology, *Pierson v. Post* actually generalizes quite easily, to encompass much of the law's view of the competition for high technology property rights.

In its decision, the court majority works through several paradigmatic questions:

1. Did Post claim a right that the law recognizes?
2. If so, by what legal principle, and at what point in time, did Post acquire[11] this right?
3. Did Post's right rise to the level of an ownership interest,[12] or was it something lesser?
4. If Post had the right that he claims, did Pierson infringe it, and when?

---

[10] *Pierson v. Post*, Sup. Ct. N.Y., 3 Cai. R. 175; 1805 N.Y. LEXIS 311 (1805).

[11] To "acquire" a right in this context means to behave, in relation to a particular resource, in such a way that the court will enforce a claim with respect to that resource. In that sense, the government confirms *ex post* that the claimant possessed the right *ex ante*, beginning with the point in time that his behavior gave rise to the right. This acquisition is really a *creation* of rights; it is a legal, not market, activity. Pierson acquired rights to the fox, but not from another market actor and not via exchange.

[12] An "interest" is the right to a benefit. An "ownership interest" is a collection of rights, generally encompassing the rights to exclude and to transfer title. Because this collection may be burdened (or "encumbered") by the rights of others, an owner's set of rights is in general a subset of all possible rights to a resource. For example, my neighbor's easement over my land burdens my ownership interest with his less-than-ownership interest. Strictly speaking, then, the "public interest" is the public's legal right to a benefit, which may burden a private resource.

This paradigm reflects the law's concern with a number of parameters that economists who wish to improve intellectual property law would do well to master: priority; possession; notice; exclusivity; term; title; prior rights; third-party rights; agency; location; jurisdiction; complementary institutions; the weight to attach to public policy and to private incentives in resolving a dispute; and constraint to resolve the controversy as it was framed by the parties, rather than as it might have been framed by a social planner. These parameters provide a useful checklist to work through, first in the context of *Pierson v. Post*, and then in the context of each of the major intellectual property regimes.

2.4.1. Priority

In this context, "priority" means "first in time." A fundamental property principle is: "First in time, first in right." Although it applies to competitive racing, priority also applies more broadly to potentially conflicting claims on a resource, even when neither party could have foreseen any "race" or conflict, and when there is only one party "racing."[13]

The critical question, particularly when the creation of property rights is in doubt, is: "First to do what?" Thus, Post was the "first to start" the fox, and first to invest resources in "hunting, chasing and pursuing the same," but Pierson was the first to *possess* it.

2.4.2. Possession

To establish priority of right to occupy a resource to the exclusion of others, one must generally demonstrate possession. "Possession" has two components: (a) the intent to possess, and (b) the capacity to control. Post demonstrated intent, but lacked capacity.[14] Pierson, having shot successfully and retrieved the carcass, demonstrated both.[15]

Investment may demonstrate intent. To the extent that its outcome is foreseeable, it may also indicate the capacity to control.[16] But, as the *Pierson* court emphasized, it is

---

[13] As I explain below, the vast majority of patent rights are not obtained as the result of any deliberate "race" against a rival, yet priority – vis-à-vis the prior art taken as a whole – nevertheless plays an essential role in the award of the rights.

[14] The court notes that it is sufficient to wound an animal mortally to establish the capacity to control it, but that to maintain intent one must not abandon the pursuit.

[15] "Intent" and "capacity" to possess interact in subtle ways. When I leave my umbrella in a restaurant, I no longer possess it, but my intent to resume possession may be inferred for some (short) period afterwards. After that point, my failure to repossess the umbrella is usually interpreted as my incapacity to repossess it, probably because I have forgotten where I left it. At that point, the law infers my "intent" to abandon it, and deems my ownership interest lost.

As will become more apparent in the extension to intellectual property, intent and capacity to possess matter in some contexts but not in others. For example, the law infers, from a trademark owner's failure to employ his mark in commerce, the owner's intent to abandon it. On the other hand, the law infers, from the general publication of a trade secret, the owner's incapacity to possess it, even if his intent to possess remains.

[16] Moral philosophers, most notably John Locke, supported one who invests his labor to claim property in the fruits of that labor. This argument from natural law provides a crucial intertemporal link between

the coalescence of the capacity to control, not the investment *per se*, that endows the investor with property rights.[17]

As between contending parties, the relative quantums of investment are generally irrelevant to determining their respective intent, capacity, or priority of rights. On the other hand, one party's *acquiescence* to the other's assertion of rights – say, by accepting a gate across a disputed path, or more generally ceasing investment in pursuit of possession – is generally accepted as evidence of intent to recognize the other's claim, and to abandon one's own.

2.4.3. Notice

Because property rights allow one party to restrict another's activity, in a dynamic setting they naturally lead to investment hold-ups. To mitigate this possibility, the law generally requires some form of notice to the public (or the adverse party) of the putative owner's rights. Thus, the *Pierson* court observes that "Pierson, *well knowing* the fox was so hunted and pursued, did, *in the sight of Post*, to prevent his catching the same, kill and carry it off." In other words, Pierson was "on notice" of Post's pursuit; similarly, Post was "on notice" of Pierson's capture. Notice serves the important signaling function of inducing investment in things that can lawfully be possessed, and away from things that are already possessed by others.

2.4.4. Exclusivity

Absent an agreement to the contrary, the right to exclude can only be exercised by a single titleholder. Thus, the fox became the property of Pierson, not of Pierson and Post jointly, even though Post's input may have been indispensable to Pierson's capture.[18] In general, the law abhors divided ownership, except by prior agreement, because of obvious hold-up problems, and because division may destroy value.[19] When legitimate competing claims exist, and division would reduce value, courts may order the

---

investment and outcome: a farmer plants because he expects to harvest. But, given the farmer's exclusive possession of the soil, planting implies future control of the harvest in a way that hunting does not.

[17] "So, also, *encompassing and securing such animals* with nets and toils, or otherwise intercepting them in such a manner as to deprive them of their natural liberty, and render escape impossible, may justly be deemed to give possession of them to those persons who, *by their industry and labor, have used such means* of apprehending them" (emphasis supplied).

[18] On the other hand, the law will enforce joint property rights when the parties demonstrate that that was their intent. For example, a marital home is a joint tenancy, from which either party may exclude all others, except that neither party may exclude the other. Parties may also create shared tenancies, in which each can exclude the other from a part of the whole. In the absence of any such agreement or pattern of behavior, however, the default rule is to assign the property rights to one of the two competing parties.

[19] Courts are more likely to divide *rights* than *things*. Remedies that divide (rights or things) are sometimes erroneously called "splitting the baby," after the Biblical story in which King Solomon directed that a live baby, claimed by each of two women, be divided in half between them. The true mother revealed herself by giving up her claim, to spare the child's life. A common, but erroneous, interpretation treats Solomon's proposal as a remedy *following* the determination of (divided) ownership, rather than as an information revelation mechanism essential to the *prior* determination of (unified) ownership. Like Solomon, courts are

resource sold and the surplus distributed between the claimants. However, because property law also recognizes that resources may have private value, and because sale and distribution eliminate private value in excess of market value, the law prefers to preserve private value and unity of ownership by, for example, facilitating the purchase of one party's ownership interest by the other.

In general, the right to exclude applies to physical "occupancy," but it does not extend to all uses by others, especially if the use is non-rivalrous. For example, Pierson could not have prevented Post from looking at his fox, or photographing it, or measuring it, as long as these uses did not interfere with Pierson's occupancy.

2.4.5. Term

In general, the right to exclude, once acquired, persists as long as the law can infer the owner's intent and capacity to control. In legal terms, Pierson's "estate" in the fox has legal, spatial and temporal dimensions: Pierson got all the rights, to all of the fox, forever. That is what "winner-take-all" means in property law. Just as courts dislike dividing a resource in law (by awarding joint exclusivity) and in space (by "splitting the baby"), they also dislike dividing it in time, absent some agreement to the contrary.[20]

2.4.6. Title

Pierson's intent, priority, possession, and notice resulted in the *creation* of his ownership interest in the fox, including the right to exclude. But, strictly speaking, his title to the fox – meaning general recognition of his legal rights – was disputed. Clear title arose only with the attempt by Post to *exercise* (what turned out to be non-existent) rights.

Because of the time lag between creation and the exercise or enforcement of rights, inefficiencies may result. Depending on the facts, the capacity to dispute title enables either party to hold up the other's prior investment and to extract rents in settlement. And of course, if title were awarded unambiguously at the moment of creation, Post would have foreseen the judgment against him, and both parties could have saved legal fees. A potentially more efficient system would award title *ex ante*, at the moment of creation, rather than *ex post*, at exercise. But such title systems are themselves costly, because they require examination of all potential title claims, not merely those subject to private dispute. For that reason, they tend to be used only with relatively costly resources, having "boundaries" that are relatively easily delineated, and which arise at a sufficient scale that the costs of administering the system can be shared by a large

---

reluctant to partition resources between claimants as a remedy, because partition may reduce the value of the resource.

[20] Agreements that convey a subset of the estate are, of course, routine. For example, a lease to a tenant conveys some of the rights (to exclude and to use, but not to transfer), for some determinate subperiod; the landlord, however, retains the residual ("forever").

number of titleholders. Well-known *ex ante* title systems cover real property (e.g., residential homes) and some forms of transportation (cars, boats, planes, etc.).

2.4.7. *Prior rights, third-party rights, and agency*

In *Pierson*, the court needed only to determine the priority of claims between the two disputants. But the court also considered the possibility of prior third-party claims to the fox. The court repeatedly refers to the fox as *feroe naturoe*, i.e., a wild (uncontrolled) animal. It is also significant that the chase began "upon a certain wild and uninhabited, unpossessed and waste land, called the beach." This "unpossessed" land is *not* "the public domain," which is land owned by the public, from which its members may not be excluded. Had the fox been, say, a cow, or had the land been owned by one or the other of the parties, or by the public or a third party, that party might have laid claim to the fox because of prior domestication, or by the principle of *ratione soli* ("according to the soil"). Moreover, Pierson's prior contractual relationships might complicate the allocation.[21]

2.4.8. *The legal role of location*

Ratione soli emphasizes the essential role that location places in the creation of property rights, including rights to resources that are not themselves attached to a place. Location is also critical to the exercise of those rights, because location often determines jurisdiction.[22]

2.4.9. *The legal role of technology*

In explicating *ratione soli*, the court distinguished the fox chase from a case in which a sportsman drew ducks to a pond on his land, through the use of decoys. *Ratione soli* gave title to the sportsman, not the shooter. Obviously, the ability to tie the object of the hunt to the "soil" depends partly on the nature of the object, and partly on the technology of hunting. Though wild ducks and foxes may each be "unpossessed," ducks are susceptible to an intermediate form of possession, not rising to the level of capture, that is difficult or impossible with wild foxes.[23]

---

[21] Under New York law, Pierson "finds for himself"; his employer acquires the fox only if Pierson had previously contracted to assign his rights to it. Other jurisdictions would deem Pierson to be an agent of a principal, even if their contract does not expressly contemplate the observed acquisition, so the fox would belong to the principal. When federal statutes create an intellectual property right (patent; copyright; federal trademarks), the statute allocates rights as between the agent and principal.

[22] It bears repeating that Pierson shot the fox in New York, thereby acquiring New York rights, which were subsequently confirmed in a New York court. For many forms of property, especially real property, any other combination of location and rights is all but unthinkable. But the centrality of location to the law of tangible property is a central *limitation* when the resource, such as information, is intangible. And as Section 5 describes, these limitations assume even greater significance across national boundaries.

[23] If, as was the custom in England, the fox had been bred (or lured for temporary capture) and then released for pursuit, the technology of fox hunting would, for legal purposes, more closely resemble the technology of duck hunting.

2.4.10. *Institutions and jurisdiction*

Property is not the only set of enforceable, non-market social rules for allocating resources. Within academic communities, for example, the rules regarding plagiarism are different from property rules, but no less enforceable. Sometimes – e.g., in evaluating conflicting claims to medical waste products – a court infers an agent's intent, hence his property interest, from the agent's conformity to practices within a particular culture (in this case, the culture of medical research using discarded body parts). The *Pierson* court, however, considered but rejected the idea that property rules should enforce the norms and practices of extra-legal institutions.[24]

Though the court has jurisdiction to decide the legal question, and to impose a remedy according to property's social rules, it lacks the authority to declare Pierson a "bad sport" or the power to impose potentially more significant social remedies, such as restricting Pierson's participation in the culture of sportsmen. If one viewed the sport of hunting as a partly cooperative, partly non-cooperative repeated game, the availability of sanctions for violating "hunting etiquette" might induce a superior investment program by all participating hunters. In short, the conclusions and remedies of property law may conflict with, and prove less efficient than, those of a specialized non-legal community.

2.4.11. *The nature of the claim*

The court is limited to deciding the question put before it. Post claimed property in the fox. The court did not decide, because Post did not claim, that Post had a separate and distinct property right to hunt without interference, or that killing the object of the hunt constitutes such interference.[25]

Pierson v. Post also illustrates the close relationship between the creation and exercise of property rights. Post's claim failed because the court determined that the right he claimed had never been created. That determination mooted any question of

---

[24] In dissent, Judge Livingston stated that the dispute

> should have been submitted to the arbitration of sportsmen. . . : they would have had no difficulty in coming to a prompt and correct conclusion. In a court thus constituted, the skin and carcass of poor Reynard would have been properly disposed of, and a precedent set, interfering with no usage or custom which the experience of ages has sanctioned, and which must be so well known to every votary of Diana. But the parties have referred the question to our judgment, and we must dispose of it as well as we can, from the partial lights we possess. . . .

The choice of forum restricts both the outcome and the range of remedies available. Indeed, the majority opinion acknowledges the court's limited power to achieve a "better solution":

> However uncourteous or unkind the conduct of Pierson towards Post, in this instance, may have been, yet this act was productive of no injury or damage for which a legal remedy can be applied.

[25] The majority opinion alludes to the existence of such a right when, in distinguishing the duck case, it observes that "there the action was for maliciously hindering and disturbing the plaintiff in the exercise and enjoyment of a private franchise. . . ."

whether or how Post might exercise his right. But suppose, instead, that Post had possessed the fox after Pierson shot it, whereupon Pierson sued Post. The court would have found that Pierson's right to exclude was created when he killed the fox, and that when Pierson attempted to exercise his right by taking possession of the fox, Post deprived Pierson of his right to exclude others. In that case, the remedy for infringement would include (a) damages to Pierson for the loss of use of the fox while in Post's possession, and (b) an injunction requiring Post to turn over the carcass to Pierson.

2.4.12. Public policy and private incentives

Property disputes are framed as a static allocation between competing private interests. But the *Pierson* dissent demonstrates familiarity with the dynamic incentive effects of property rights, and their social consequences. In particular, the dissent worries about private underinvestment in socially beneficial activity, and – contrary to the Coase Theorem – that the equilibrium dynamic outcome is not independent of the initial allocation of rights.[26] However central is this policy concern to promoting dynamic efficiency, it lacks an essential element of the common law of property: the representation of a private interest.[27] Moreover, as the majority opinion points out, there may also be efficiency consequences to the dissent's proposed reassignment of rights.[28]

The court did not say – because it is irrelevant to the law of property – whether foxes are hunted more efficiently when title is awarded to the pursuer or the shooter. The absence of any legally cognizable mechanism with which to evaluate alternative investment incentives may not be surprising in the context of chattel or real property. But it is perhaps more noteworthy that the exact same gap appears in intellectual property systems, which derive their legal justification, at least in part, from their alleged

---

26

> By the pleadings it is admitted that a fox is a "wild and noxious beast." ... His depredations on farmers and on barnyards, have not been forgotten; and to put him to death wherever found, is allowed to be meritorious, and of public benefit. Hence it follows, that our decision should have in view the greatest possible encouragement to the destruction of an animal.... But who would keep a pack of hounds; or ... pursue the windings of this wily quadruped, if ... a saucy intruder, who had not shared in the honors or labors of the chase, were permitted to come in at the death, and bear away in triumph the object of pursuit?

27 Hunting foxes may be socially valuable, but "society" has no private right to have foxes hunted (nor, for that matter, to prevent their being over-hunted). Like a donut, property systems can be defined by the hole at their center. In this case, the "hole" is the public's lack of standing to assert any claim to property in foxes, fox-free zones, or fox-hunting resources.

28 "If the first [person] seeing, starting or pursuing such animals, without having so wounded, circumvented or ensnared them, so as to deprive them of their natural liberty, and subject them to the control of their pursuer, should afford the basis of actions against others for intercepting and killing them, it would prove a fertile source of quarrels and litigation."

capacity to stimulate investment.[29] In short, the ends of intellectual property are rooted in dynamic economics, but the means are rooted in the law of tangible property.[30]

3. *Intellectual property: patents*

An economist who understands what Pierson got, why he got it, and why the problem is framed as: "Who has property in the fox?", has a surprisingly broad foundation for the "law and economics" of intellectual property. But intellectual property law also varies, not only from the law of tangible property (Section 2), but also between types of property (Section 4) and across jurisdictions (Section 5). From a "law and economics perspective," some of these differences have greater significance for economists than they do for lawyers, so it is important to highlight them.

To make concrete the comparison between tangible and intellectual property, I focus initially on the differences introduced by the patent system. Among the various intellectual property systems, the legal rationale for patent system – the "progress of the useful arts" – is most closely identified with the economic rationale for market intervention: to promote investment by competing producers of a non-excludable good, to stimulate productivity growth by users of a non-rivalrous good. The patent system serves, in turn, as the basis against which to evaluate other intellectual property systems.

---

[29] As the facts in *Pierson* demonstrate, mere investment is not, by itself, either necessary or sufficient to create title to property. But even if investment were relevant to determining a rights winner, intellectual property law lacks any jurisdiction over the fundamental *policy* question: is *total* investment (among all competitors, before and after the award is made) *socially* optimal?

[30] This is not to say that the law of tangible property is immutable, or that subtle factual differences cannot induce markedly different outcomes that suggest ways to reform intellectual property law.

Example. In a recent case, a U.S. judge was asked to decide the ownership of the baseball that Barry Bonds hit for his record 73rd home run of the 2001 season. The contending parties were a first man who reached for and touched the ball, but who was knocked down by the crowd, and a second man who eventually took possession of it. Apart from complications unrelated to *Pierson v. Post* (e.g., who owned the baseball when it left the pitcher's hand? was it abandoned?), the case was notable because the court found that, unlike Post, the first man had obtained a "pre-possessory interest" in the ball, by dint of his thwarted pursuit, equal in magnitude to the interest of the eventual finder. (It was critical, in this case, that interference arose, not from the finder, but exogenously from the crowd.) Because of the parties' joint and adverse ownership interests, the court ordered the ball sold and the proceeds divided between them. *Popov v. Hayashi* (S.F. Sup. Ct. CA, 2002). Note that the court did not "split the baby" (ball), which would have destroyed its value.

Although the case raises the intriguing possibility, for intellectual property purposes, that an unsuccessful pursuer might acquire some rights to the object of pursuit, it also illustrates the subtle differences in the role that property rights play in stimulating investment: a baseball, once hit and retrieved, has realized its entire market value, but an invention or other information good usually remains commercially undeveloped, its full value only being realized after further specific investment by one of the disputing parties. In such cases, even the sale of the "fox" to a third party (and division of the proceeds among the pursuer(s)) diminishes or destroys its value.

3.1. The patent system and property law: parallels

On a first pass, it is easy enough to analogize *Pierson v. Post* to a winner-take-all patent race (see Reinganum, 1981 and subsequent refinements). An inventor, being first in time by definition, is also first in right, i.e., he obtains *priority* over all others. Here "first" means "first to 'reduce the invention to practice,' " i.e., to *possess* or control the invention. Like Pierson's fox, the invention is drawn from a previously "unpossessed" resource, to which no one has *prior rights*.

As a condition of the patent right, the inventor must provide *notice* to the public, "particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention."[31] *Title* to the patent estate encompasses several property rights that are jointly created, including the rights to *exclude* all others, from making, using or selling the invention, and which the patentee may enforce, jointly or severally, for periods up to and including the maximum *term*.

As with hunting, the *location* of the act of invention matters, both in geographical and technological spaces. Inventions that occur outside the patent *jurisdiction* may be subject to different standards of proof than those occurring inside, so that "possession" means something different abroad than it does at home. Similarly, and as I explain in greater detail below, an invention that is captured on the technological "soil" previously claimed by another may be subject to the other's prior rights *ratione soli*.

Conversely, the *technology* of invention – meaning the manner in which the inventor achieves possession – is mostly irrelevant.[32] As Section 3.2 explains, there are subtleties to the definition of "inventing" that complicate the determination of what – exactly – constitutes "possession." As a practical matter, an inventor's ability to appropriate returns on his invention, and even to define the boundaries of an invention, varies by technology field.[33] Even if it did not, national patent laws sometimes restrict the technology fields that can be subject to patent protection; as Section 5.2 explains,

---

[31] *See* 35 U.S.C. §112; other jurisdictions are similar. Separately, the law also requires that the patentee cannot collect damages until he provides notice of infringement, either by marking his product with a patent number, or by filing suit against the infringer. *See* 35 U.S.C. §287.

[32] 35 U.S.C. §103: "*Whoever* invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor. . . ." (emphasis supplied). Thus (contrary to the court's remark in *Pierson v. Post*), it would be wrong to award an invention to one who had "laid the groundwork" for it in "the soil" (e.g., attracting ducks to a pond). The inventor must be a shooter. Potentially fatal complications ensue, however, when inventors are omitted from the patent, or non-inventors are credited. In short, patent law is neutral to the manner in which the invention is produced, but it is not indifferent to the identity of the producers.

[33] For example, patents are important in the pharmaceutical sector for several technology-related reasons:

1. *Low imitation costs*. Once its molecular structure is disclosed, an active chemical ingredient usually can be copied exactly at low cost.
2. *Clear invention boundaries*. Pharmaceutical inventions generally rely on a single patent to protect a single active ingredient. Because of the well-known rules of chemical structure and interaction, the patent's claims can be drafted with great precision; for the same reason, infringing compounds can also be identified with great accuracy.

however, recent innovations in international institutions have reduced this variability across countries.

The role played by non-property *institutions* in supplementing intellectual property rights, and their legal (or extra–legal) authority over the inventor and his rivals, lie well beyond the scope of this paper. But in analogizing to *Pierson*, it suffices to observe that the club of research "sportsmen" has not stood by idly while the courts award foxes to saucy intruders.[34] For their part, as Section 5 explains, legislatures have effectively broadened the economic jurisdiction of the courts to include foreign activities.

3.2. The patent system and property law: differences

The breakdowns in the analogy between patent systems and tangible property systems occur in myriad ways. I focus on those that are most relevant for economists. The first and most important is: "What is the fox?" For economists, this turns out to be a much harder question than it first appears.

3.2.1. "The fox" and protectible subject matter

Because patents are "intellectual" property, and because they typically arise following some kind of search process, it is natural to view the "fox" as new information, an otherwise non-rivalrous and non-excludable good that the patent system induces inventors to find and permits them to capture. But this view is legally and economically erroneous.

An invention is "any new and useful process, machine, manufacture, or composition of matter. . . ." A patent creates rights in physical, tangible things: "useful" means, in patent law, "capable of industrial application." Perhaps more importantly, the exercise of these rights excludes only physical actions occurring in a physical place. These things and actions necessarily require new information to be made or performed ("any

---

3. *High invent-around costs.* Because of the complex interaction between an active ingredient and human physiology, and the relatively high regulatory standards for safety and efficacy, the cost of inventing around (discovering a non-infringing alternative having equivalent therapeutic efficacy) often approaches or exceeds the cost of the initial discovery.
4. *Coordinated development.* By controlling the patent on the active ingredient, the patentee usually controls all of the important intellectual property that is needed to produce it, as well as all improvements (e.g., extended release ("one-a-day") formulations). Therefore, the fruits of successful research are usually not diluted by licensing payments or hold-up problems.

[34] In the United States, for example, universities grew tired of pursuing basic research results, only to see the practical applications of those results patented by others. Rather than change the patent system, however, they changed their agency relationship: whereas they once "found" inventions on behalf of the research funder (typically the federal government, which retained title and a license to the intellectual property), under the University and Small Business Patent Procedures Act of 1980 ("Bayh–Dole Act") they now receive title themselves. This change in ownership has sharply increased the rate of university patenting; some have argued that it has tilted universities towards applied research and away from basic research.

new and useful. . ."), but they are not themselves information. In other words, the subject matter to which the rights relate – the fox – is new information "embodied" in a properly described physical thing or process – not the new information itself.[35]

3.2.2. Notice and other signaling

This distinction matters greatly to an economic interpretation of intellectual property law, and of patents in particular. A patent has two parts: the specification, and the claims. The claims declare to all others what the invention is, analogous to a deed to real property. These fulfill property law's *notice* requirement, and they determine the scope of the right. But the patentee must disclose additional information in the specification, including how to make the invention. By tying new information to a physical manifestation, the patent system induces investment in an excludable good (the claimed embodiment) that produces, as a necessary byproduct, a non-excludable good (the disclosed information). In short, "notice" in patent law assumes a strange (to classical property law) combination of traditional and non-traditional signals: "this is my fox; here's how you catch one."

Because the pure information disclosed by a patent remains non-rivalrous, it is subject to all of the same conceptual difficulties that information possesses in other contexts. Probably the most important of these difficulties, at least for productivity-oriented economists, is capturing the spillover effects on rivals' production and search functions. Although many economists have recognized this problem, they generally

---

[35] A useful paradigm for this distinction is found in the U.S. Supreme Court's decision in *O'Reilly v. Morse*. In addition to the telegraph itself, Samuel Morse claimed

> the use of the motive power of . . . electro-magnetism, *however developed*. . . , being a[ny] new application of that power of which I claim to be the first inventor. . . (emphasis supplied).

While enjoining O'Reilly's infringement of Morse's telegraph claims, the Court invalidated his electromagnetism claim:

> The [patent] act . . . requires that the invention shall be so described, that a person skilled in the science to which it appertains . . . shall be able to construct the improvement from the description given by the inventor.
> Now, in this case, there is no description but one, of a process by which signs or letters may be printed at a distance. And yet [Morse] claims . . . a patent, for an effect produced by *the use of electro-magnetism distinct from the process or machinery necessary to produce it*. The words of the [patent act] show that no patent can lawfully issue upon such a claim. For he claims what he has not described in the manner required by law (emphasis supplied).

56 U.S. 62 (1853)

Although the contrast here is particularly stark – Morse disclosed pure information (the existence and properties of electromagnetism) but he claimed an invention (the telegraph) – it is generically the case that an inventor does not control the information he discloses, but only the invention.

do not specify it precisely.[36] For their part, intellectual property lawyers also usually misstate the legal role of information disclosure.[37]

3.2.3. Possession and priority

Under the principles of property law, it is so important to distinguish that which can be possessed from that which cannot that the law refuses to grant patents on pure information. The law's most important conceptual distinction in this regard is between invention and the discovery of natural or scientific laws or principles (such as electromagnetism). The latter, being in the public domain by definition, cannot be possessed privately. Invention, on the other hand, is a human activity resulting in a human creation.[38]

The determination of *priority* – i.e., the point in the pursuit when the law declares a winner and the race is over – has been subject to different interpretations, both across jurisdictions and over time. Most of the world outside the United States adopts a simple "first to file" rule regarding patentable inventions: "possession" is indicated by filing an acceptable application, so whoever files first in time is first in right. This rule is strictly analogous to that adopted by the *Pierson* court: possession is what matters; pursuit is immaterial. On the other hand, U.S. patent practice defines an inventor to be the first person who conceives of an invention and who subsequently (not necessarily first) possesses it, by reducing the invention to practice (either by actually making it or

---

[36] What matters to an analysis of the role of information disclosure is the incremental disclosure contained in the patent's specification, beyond the information revealed by (a) publishing the claims themselves, and (b) the inventor's simply using or selling the invention. The "enabling disclosure" requirement attempts to standardize the level of information disclosure, thereby reducing the variability that would otherwise occur because of differences in technology (processes reveal less information about the invention than products), or in the nature of commercial exploitation (inventions that are not marketed reveal less than those that are).

[37] The usual legal justification for mandating disclosure is the "contract" between an inventor and society, under which the inventor receives an exclusive right in exchange for his description of how to make the invention. In the absence of disclosure, so the argument goes, the contract would fail for lack of consideration. (In lay terms, a contract is invalid unless the party receiving the promise gives up something in return.)
Of course, there is no "contract" between an inventor and society; this analysis is merely metaphorical. But even if it were not, it is factually incorrect. In addition to whatever private value is captured by the inventor, economists would include, in the social value of the invention: (a) increased consumer surplus; (b) reduced imitation cost (by shifting rivals' current production function), resulting from the publication of the claims and the use/sale of the invention; (c) reduced search costs (by shifting rivals' R&D function), by the same mechanisms; and (at the expiration of the patent) (d) additional reductions in deadweight loss and input costs, from competition. To the extent that social value exceeds private value, the inventor necessarily provides consideration to society. The level of disclosure is simply a "contractual" parameter that can be manipulated to vary the distribution of gains between the inventor and the public.

[38] Over time, the legal boundary between discovery and invention has been systematically pushed back, in favor of invention, such that previously unpatentable material – genetically modified plants and microorganisms, or mathematical algorithms embedded in computer programs – are now treated as inventions in many jurisdictions. Much of the policy argument over computer software patents and gene patents derives from the conceptual struggle over what constitutes pure information and what constitutes an industrial application of that information.

by filing an acceptable application). Under the American rule, the person who is first to conceive, but second to file, establishes priority.[39]

Because an invention must be "new" and "useful," a putative inventor faces a fundamentally different tradeoff than did Post when he pursued the fox. By ending the chase too early, an inventor may discover that he has partially or entirely failed to possess an invention, because: (a) he has merely made a discovery, not an industrial application; (b) he does not know how to tell others how to reproduce the invention for themselves; or (c) he does not know how to claim the invention in such a way that others cannot easily imitate it. On the other hand, the longer the inventor pursues the invention, the better to secure his possession, the more likely that others' information disclosures will render the invention "non-new" when possessed.

3.2.4. Exclusivity

Probably the most important departure from property law, at least from the perspective of traditional competitive analysis, is found in the bundle of rights that a patentee receives. Although a patentee receives the paramount property right – the right to exclude others from the fox – he does not receive the right to use. The competitive consequences of this distinction are difficult to overstate.

To clarify this departure from the law of tangible property, it is helpful to think of property rights in a car. With title comes the key to the car. The key enables the owner to exercise two distinct property rights: the right to exclude others (by locking the door), and the right to use (by starting the engine). In patent law, these are two separate keys, but the patent holder only receives the locking key. He may well have to negotiate with someone else for the ignition key. Further, he may not know at the

---

[39] The American rule is exactly analogous to the more general "mortal wounding" rule that the *Pierson* majority cited (but did not apply, because Post's pursuit did not include mortal wounding):

> ... actual bodily seizure is not indispensable to acquire right to, or possession of, wild beasts; but that, on the contrary, the mortal wounding of such beasts, by one not abandoning his pursuit, may, with the utmost propriety, be deemed possession of him; since thereby the pursuer manifests an unequivocal intention of appropriating the animal to his individual use, has deprived him of his natural liberty, and brought him within his certain control.

In other words, neither starting the fox nor imagining an invention is sufficient to create property. The pursuit must end in possession, but as between two successful pursuers, the U.S. interpretation of the "first in time, first in right" principle assigns priority to the one who conceives in detail ("wounds") first and who demonstrates "diligence" – i.e., "not abandoning his pursuit" – in reducing to practice.

Similarly, much of the expansion of "invention" into the realm of "discovery" has had the effect, intentionally or not, of pushing back the definition of capture to earlier and earlier points in the chase. As one might expect, this expansion has been sought by those, like universities and privately funded spinoffs from university labs, that have a comparative advantage in the early stages of pursuit.

Finally, this example also illustrates the importance of national legal procedure in establishing international intellectual property practices: because the U.S. constitution authorizes Congress to secure rights for "inventors," and "inventor" means "first to conceive plus diligent reduction to practice," most commentators agree that the United States could not adopt the "first to file" system used in the rest of the world without amending its constitution.

time of purchase whether he has received a "one-right key" or a "two-right key." One can easily imagine the complexity that this property arrangement would introduce into both buyers' and sellers' decisions: hold-up problems and other forms of strategic behavior would increase the costs of a transaction that both automobile consumers and economists regard as routine. And the consequences for competitive equilibrium are also not trivial: clearly, the market price for otherwise identical cars would depend on the property rights that accompanied the car. That fact would, in turn, influence production decisions and perhaps create secondary markets in "automobile use rights."

This bizarre and apparently inefficient market outcome, which results simply from removing the right to use from the owner's bundle of rights, is more or less familiar to economists who try to model the exercise of patent rights. The reason is that the government may grant a patent on an invention – say, an improvement to a prior invention – which, if practiced, would constitute infringement of the prior patent. Because a license is a contract that conveys the right not to be excluded (i.e., the right to use), secondary license transactions may be quite significant, not only for non-patenting imitators who wish to use the original invention, but also for patent-holding "improvers" contracting for the right to use their improvements.[40]

As in the case of the fox, there are limitations on a patentee's right to exclude others from non-rivalrous activities. Two are particularly important to competition and productivity analysis. First, "exhaustion" deems a patentee's rights to have been fully exercised after his first sale, and prevents him from acting against subsequent users or resellers. Exhaustion limits a patentee's ability to price-discriminate. Second, the "research exemption" allows others to make the invention for certain research purposes (e.g., a generic drug manufacturer may produce a patented composition in the course of satisfying the regulatory approval process for new drugs[41]), but not others (e.g., the "academic" use of a patented research tool in projects that may have commercial application[42]). Curiously, actually making the invention (for the purpose of verifying that the inventor has disclosed enough information to enable others to make it without "undue experimentation") is not one of the exemptions from infringement.

3.2.5. Prior rights and the principle of ratione soli

The enhancements to the patent law's notice, priority and exclusivity requirements jointly imply a dynamic property relationship among successive inventions unlike anything found in the law of tangible property. When Pierson captured the fox and the court recognized his right to it, the main resource consequence was one less fox for everyone else. But there was little about *fox-hunting* that had changed.

---

[40] Even though licensing is widespread, however, it is generally erroneous to hypothesize a "license market," because any given right to use generally can be obtained from exactly one firm. Even when there exist close substitutes in the product market, it is unusual to find close substitutes for the right to use a patent property: either there must be multiple entities who can grant the right to use, or if the invention is an improvement, it must improve at least two patented inventions that are themselves substitutes.

[41] *Merck Kgaa v. Integra Lifesciences I, Ltd., et al.* No. 03-1237 (Sup. Ct. 2005).

[42] *Madey v. Duke University*, 307 R.3d 1351 (Fed. Cir. 2002).

On the other hand, when an inventor obtains a patent, he adds to three stocks: (a) by establishing first possession, he adds a new and useful productivity improvement,[43] from among previously unidentified improvements, to the current stock of production technology; (b) by notifying others how they may reproduce his possession, he expands the public domain – the stock of information that is "owned" by everyone (i.e., from which no one can be excluded); and (c) by enumerating his claims, he expands the stock of private rights, and thereby increases restrictions upon others, not only in their use of the patented invention, but also in their pursuit of other inventions.

Following the first invention, a patentee's increased profitability induces his rivals to search for a previously unpossessed substitute, using the expanded public information set, which (all else equal) steers them directly towards the rights held by the patentee. By the principle of *ratione soli*, these rights represent the expanded territory from which a subsequent invention is likely to be drawn: the (second) fox may be unpossessed, but the land on which a rival pursues it is not. Thus, the patent system's overall inducement is a complex mixture of reduced and increased costs for subsequent inventors.[44]

In short, the two parts of a patent have opposite effects: the information disclosed in the specification remains non-rivalrous, while the claims are "super-rivalrous." Economists sometimes refer to this phenomenon in the aggregate as "cumulative innovation" (Scotchmer, 1991), but the actual degree of rivalry for the initial productivity improvement depends on the interaction of the other two rivalries. When super-rivalry dominates, one outcome is a "patent thicket": overlapping patent rights that prevent entry into the product market (Shapiro, 2001).

Ironically, commentators fault the patent system *both* for its absence of complete property rights (in particular, the right to use), *and* for its supposedly "winner-take-all" nature. For example, most economic models of innovation postulate a first ("pioneering") inventor, whose rights are unrestricted, and second ("follow-on") inventor, whose rights are restricted by the first inventor. Yet these models lack convincing grounds for differentiating between a pioneering invention and the patented improvements to it. As the number of patent lawsuits directed at successful, market-defining

---

[43] Here, "improvement" must be interpreted legally, not economically. An improvement may be legally patentable even if it is productivity-reducing in all states of the world (except (perhaps) on a set of measure zero).

[44] Following behind the first inventor, subsequent hunters face different incentives and opportunities as a result. As with foxes, a successful hunt reduces the stock of productivity improvements available; that makes hunting harder. The first inventor gains a cost or quality advantage in the product market; that makes hunting more profitable (or, equivalently, the failure to hunt is more costly). There is additional public information about how to hunt, and about the location (in product and/or technology space) where the first hunter thinks hunting is profitable; that makes hunting cheaper. This additional public information also raises the bar for patentability; that makes hunting more expensive. And the first hunter receives additional rights that restrict subsequent hunters' ability to exploit their inventions; that makes "nearby" hunting less profitable, but "distant" hunting more profitable. Whether the net effect of these opposing influences actually results in "progress" is anybody's guess, although there is substantial indirect evidence that the patent system increases productivity growth in the aggregate (Putnam and Tepperman, 2005).

products shows, even pioneers build on prior inventions, which they too lack the right to use. In the absence of that right, a patent race winner cannot really "take all." Put differently (and contrary to most models), nearly *every* inventor is, potentially, a "second inventor."[45]

Needless to say, the patent system's omission of the right to use is critical to the analysis of competition, because it creates vertical relationships among horizontal competitors. Both the contractual resolution, as well as the absence of resolution, of these relationships have raised antitrust concerns.[46]

3.2.6. Title

Unlike Pierson, the patentee receives *title* to his invention via an *ex ante* examination by the patent office. But examination is imperfect, which may impair title until the patentee attempts to enforce his rights. The uncertainty of patent examination is so high that economists have begun to refer to "probabilistic" property rights (Lemley and Shapiro, 2005).[47] The systematic uncertainty surrounding the scope and validity of issued patents has led at least one commentator to claim that the patent office is, and should be, "rationally ignorant" of patentability.[48]

---

[45] The probability that the practice of one's invention infringes another's patent depends in general on the field of technology, and in particular on the interdependence of product inputs. In traditional pharmaceuticals, a patented chemical compound is relatively unlikely to infringe a prior patent, unless it is member of a class of compounds that was previously claimed but that exhibits some unforeseen attribute. On the other hand, among biotechnology-based pharmaceuticals, the use of patented research tools and other inputs means that an inventor is relatively likely to require a license from a prior inventor. Among complex electronic devices, such as microprocessors, the probability of overlap is very high, which is the main reason why microprocessor manufacturers and users maintain a web of cross-licenses to their entire patent portfolios.

[46] When firms pool intellectual property, the contract may be challenged as a form of horizontal price-fixing. See *In the Matter of Summit Technology and VISX*, FTC Docket No. 9286 (1998). On the other hand, when one firm rejects potentially welfare-improving cooperation in the use of its intellectual property, its rejection can be characterized as a "refusal to deal," i.e., an attempted monopolization. *In re ISOs Antitrust Litigation*, 203 F.3d 1322 (Fed. Cir. 2000). In the United States, the antitrust agencies generally have acknowledged the right of an intellectual property owner to refuse to deal with competitors, at least if the refusal is unconditional (U.S. DOJ-FTC, 2007), but in Europe the view is more mixed.

[47] As a quasi-contract, a U.S. patent's claim scope is a question of law, to be construed by a judge. Despite the best efforts of the patentee and the patent office to endow the claims with precision, litigation often reveals a fair amount of elasticity in the interpretation: the patentee wants a broader construction, which aids in proving infringement, but not so broad that the claims can be invalidated by prior art; the accused infringer generally prefers the opposite. Given the judge's construction, the jury determines invalidity and infringement, which are questions of fact.

[48] *See* Lemley (2001). It is more efficient, so the argument goes, to let private parties bear the costs, via subsequent litigation, of determining the ultimate scope and validity of that small subset of patent rights that are actually disputed. This view ignores the substantial investment externalities that arise when the true scope of patent rights is unknown (Putnam and Tepperman, 2004). One of the patent office's most important contributions is to internalize these externalities, by ensuring the accuracy of patent claims and disclosures.

3.2.7. Location

Most of the interesting law-and-economics issues associated with location arise when crossing national boundaries. Since that is the subject of the next section, I again defer the bulk of the discussion. But even when rights are created uniformly within one jurisdiction, local enforcement practices may substantially determine their exercise.[49]

3.2.8. Technology

Here "technology" means "the manner of possessing the invention," not the invention's field of application. For example, at one time U.S. courts imposed a "flash of genius" test on inventors. Since most invention is not accompanied by a "Eureka!"

For example, in a recent Canadian case, the Federal Court of Appeal upheld a "speculative" (at the time of filing) patent, on the grounds that the patentee was only required to prove its utility at trial. The Supreme Court of Canada reversed unanimously, recognizing the economic value in a clear public record:

> The doctrine of "sound prediction" balances the public interest in early disclosure of new and useful inventions, even before their utility has been fully verified by tests, and the public interest in avoiding cluttering the public domain [*sic*] with useless patents and granting monopoly rights in exchange for speculation or misinformation.

Apotex Inc. v. Wellcome Foundation Ltd., 2002 Sup. Ct. Can. 77.

[49] Enforcement requires its own complex litigation and licensing strategy, a (small) subset of which is location-dependent:

1. *Choice of district court.* In recent years, the sparsely populated Eastern District of Texas has taken a disproportionate role (almost 10 percent of the more than 2700 patent suits filed in the United States in 2006) by promising expedited litigation. This promise favors parties (typically plaintiffs) who assemble their case prior to filing. As long as non-trivial infringement occurs within the district, the patentee's choice of venue is proper. This provision captures most patent-intensive consumer goods like computers and pharmaceuticals.
2. *Choice of appellate court.* Prior to the creation of the Court of Appeals for the Federal Circuit (CAFC) in 1982, U.S. patent appeals were heard by the eleven regional circuit courts. Because the Supreme Court (which resolves conflicts among the circuits) rarely accepted patent appeals, conflicting interpretations of the law persisted. This geographic variation produced "forum shopping" – endogenous selection of a favorable circuit – and additional uncertainty about the true scope of the patent. By unifying all appeals in a single court, the CAFC – which is often associated with a systematic "strengthening" of patent rights – reduced regional variation and uncertainty, the economic effect of which may have been equivalent to "strengthening," even if the only regime change amounted to increased uniformity across circuits.
3. *Choice of adversary.* Courts, and juries in particular, may be biased towards domestic firms; property principles may magnify that bias. In *Graham v. John Deere*, 38 U.S. 1 (1966), the U.S. Supreme Court developed certain secondary indicators of a patent's validity, including "acquiescence by others" (see also footnote 106). Under this principle, U.S.-based patentees find it optimal to sue foreign firms, which are adverse to litigation if domestic bias exists, and which are therefore more likely to "acquiesce" to the asserted right by settling the litigation. The patentee can then cite this acquiescence in support of the patent's validity when litigating against other U.S. firms.
4. While these location-based strategies might seem highly idiosyncratic to the United States, they are largely mirrored in Europe, where patent rights are created by a single regional authority (the European Patent Office), but interpreted and enforced by national courts.

moment, this standard for possession invalidated many otherwise non-obvious inventions. The pendulum then swung the other way, with the CAFC upholding inventions as long as there was no explicit motivation in the prior art for combining other prior art elements as the inventor did. Recently, the U.S. Supreme Court began to undo this "narrow, rigid" rule, in favor of a more "flexible" (if less transparent) rule that raises the bar on the standard of possession that the inventor must show.[50] The Court also acknowledged that the manner of possession, and therefore the proper application of the tests for obviousness, may differ across fields of application.[51]

3.2.9. Term

Commentators sometimes distinguish patents from "true property" because the ownership interest is not permanent. But this distinction turns out to be weak. The U.S. constitution refers to "exclusive rights" for "limited times," not to "property" *per se*, but the statute states otherwise.[52] Accepting that patents are property, even tangible property rights (including real property rights) may also be impermanent: a 99-year lease is no less "property" because of its finite term. In any event, other types of intellectual property, notably trademarks and trade secrets, have no statutory limit to their terms.

3.2.10. Public policy and private incentives

The genius of intellectual property systems is that they are decentralized: inventors largely define whether and what to patent, it is sufficient for the government to establish "boundaries" and award title. The creation of intellectual property rights requires relatively little information, and in particular little reliance on expectations or other economic decision-making. Similarly, the enforcement of intellectual property rights – that is, the determination of liability and an injunction – is also almost entirely free of economic input. Atomistic optimization *follows* from the sub-atomic rules that govern the award of property rights. This economy of decision resources should not go

---

[50] The court explicitly recognized that, when the standard of possession is too lax, the resulting equilibrium may reduce productivity growth:

> Granting patent protection to advances that would occur in the ordinary course without real innovation retards progress and may, in the case of patents combining previously known elements, deprive prior inventions of their value or utility.

KSR International Co. v. Teleflex Inc., 550 U.S. _ , 127 S. Ct. 1727 (2007).

[51]

> The diversity of inventive pursuits and of modern technology counsels against limiting the analysis in this way. In many fields it may be that there is little discussion of obvious techniques or combinations, and it often may be the case that market demand, rather than scientific literature, will drive design trends.

Id.

[52] "Subject to the provisions of this title [35], patents shall have the attributes of personal property." 35 U.S.C. §261.

unremarked, because it is even more pronounced than is the contrast between the information required for decentralized competitive equilibrium and that required by a central planner.

Yet the decentralization of intellectual property, and its independence from economic decision-making, are also sources of maddening indeterminacy. Unlike antitrust laws, which explicitly take into account (some version of) consumer welfare, intellectual property laws do not endow the public with a distinct, legally cognizable interest, even though the chief justification for intellectual property is the "public interest." The public has no *property*[53] interest to infringe.[54]

In the past fifteen years, issues at the interface of antitrust and intellectual property rights have occupied an increasing share of antitrust agencies' analytical and enforcement efforts (U.S. DOJ-FTC, 1995, 2007). Some have argued, explicitly or implicitly, that antitrust enforcement naturally fills the vacuum created by the absence of a well-defined ownership interest held by the public. Yet the tools of antitrust have their own limitations as a counterweight to private property. Regulators are fond of observing that, "Antitrust protects competition, not competitors."[55] But the public's underlying claim within intellectual property systems is to *progress*, not to competition *per se*.[56] Simply stated, antitrust law protects competition, not progress.

---

[53] The alleged existence of a true public interest was the subject of a challenge to the U.S. Copyright Term Extension Act of 1998. CTEA changed the copyright term following an author's death from 50 to 70 years (to conform to European practice). The controversy concerned the application of this extension to already existing works, not new works. The law was alleged to be unconstitutional *per se*, because it did not, as Art. 1, §8, cl. 8 requires, "promote the progress of science. . . ." Since existing works had already been induced by whatever incremental incentive the "life-plus-50" term provided, extending their term of protection could only delay their entry into the public domain, without any countervailing benefit to the public as to these works. In other words, the public's interest in "progress" was unambiguously harmed, rather than promoted. Moreover, by bestowing an additional windfall gain on existing works, Congress had failed to extract any *quid pro quo* that is often said to be necessary to the intellectual property "contract." In *Eldred v. Ashcroft*, (01-618) 537 U.S. 186 (2003), the Supreme Court rejected these arguments, basically deferring to Congress to determine whether CTEA promoted progress as a whole, rather than with respect to any subset of works.

[54] To be sure, intellectual property rights can be misused, and this is sometimes couched as an encroachment upon the public domain. For example, courts have declared unenforceable that portion of a patent license extending beyond the term of the patent, on the grounds that the patentee is attempting illegally to extend his estate in time, though the *per se* illegality of this practice within antitrust circles is fading (U.S. DOJ-FTC, 2007). But even in such cases, the licensee does not stand in the shoes of the public; the public domain and the public interest are not the same. For example, contrary to the licensee, the public might well prefer a license that overhung the end of the patent's term, thereby encroaching on "the public domain," if the alternative was that (a) the patentee did not license at all and (b) other competitors could not or would not enter when the patent expired.

[55] "The goal of competition policy should be to protect competition – not competitors" Masoudi (2006).

[56] Competition is a poor proxy for progress for many reasons:

1. The patent system, like other intellectual property systems, is itself an exception to antitrust's unambiguous preference for competition. It is simply an article of faith within antitrust law that the patent system is welfare-increasing. Nothing in antitrust law compels, or even suggests how to test the truth of, that proposition. Just as current antitrust laws could not generate a patent system in theory, or evaluate

3.2.11. *Patent racing*

Although a fox-hunt is a helpful metaphor for learning about the *creation* of property rights, particularly in the context of a race, it is important not to overstate the importance of "racing." When Post goes to hunt another fox, he will probably just bag one without interference. The Pierson incident is a low-probability event, however useful it is for defining rights. Even more than foxes, most patents are not captured after a well-defined two-person race.[57] Creativity being what it is, competitors are rarely "racing" for the same thing. Foxes are already defined, whereas inventors possess much greater control over what constitutes "an invention." Even if competitors agreed on a desirable new product or set of product characteristics (i.e., what an economist would consider "an invention"), this larger product construct often encompasses many smaller legal inventions. As a matter of strategy, it is usually worthwhile patenting intermediate research results or inputs, thereby retaining some control over the final product. The

---

a patent system empirically, so they cannot generate or evaluate a proposed modification of the patent system that alters the relationship between dynamic incentives and static competition.

2. When intellectual property protects quality improvements, it is impossible to characterize "progress" in economic terms without reference to quality-adjusted prices. Neither antitrust jurisprudence nor antitrust economics pays more than lip service to quality-adjusted prices when defining "the competitive price level" and related market benchmarks.
3. Competition in intellectual property also takes non-market forms that run counter to both traditional antitrust analysis and traditional property analysis. For example, a pursuer who is behind in a patent race is legally permitted – indeed, encouraged – to interfere with the leader's pursuit by disclosing information that renders the leader's capture "insufficiently new."
4. There is as yet no formula, analogous to the "Hand formula" in tort law (which implies a duty of care when the cost of care is less than the expected loss from the absence of care), for determining whether a given exercise of intellectual property rights would justify a finding of liability, and no simple one appears to be in the offing. But the deeper legal and economic question (to which neither antitrust nor intellectual property law has an answer) is: "Must *every* intellectual property right individually promote progress,' or is the promotion of progress merely a necessary condition of the system as a whole?" For example, an optimal intellectual property policy might nevertheless give rise, with some probability, to competitively inefficient monopolies. Suboptimal antitrust reform may eliminate statically inefficient monopolies as they arose, in the process also eliminating the dynamic competition for them that made the original policy stochastically optimal. Whether or not that proposition is true, the larger point is that antitrust law lacks both the economic tools and the legal jurisdiction to determine its truth.

[57] Some commentators claim that racing is uncommon, based for example on the low incidence of "interferences" in U.S. patent prosecution. The patent office declares an interference when two or more pending applications claim the same invention. The office then conducts a hearing to determine which of the contenders was the first inventor, based on the "first to conceive followed by reduction to practice" definition. But counting interferences probably understate the true incidence of racing, because the publication of an application, or even an announcement of it, causes other competitors to drop out without filing a second application.

In the vast majority of cases, the first to conceive of an invention is also the first to file an application on it. This observation has led many to observe that the United States could join the rest of the world in moving to a "first-to-file" definition of inventorship with little change in either private ownership or public exploitation of inventions. But the move has been opposed by small inventors, who believe they have a comparative advantage in conception but a comparative disadvantage in reduction to practice, relative to large firms.

result is that the first introduction of a new product is rarely co-incident with winning all of the intermediate "races" for the constituent intellectual property inputs.

For example, an important product innovation in the U.S. automobile market was the minivan, which was "invented" by Chrysler in the 1980s (Petrin, 2002). Initially, Chrysler's minivan occupied a unique niche in the automobile ecosystem: for example, it had greater cargo space than a station wagon, but it was built on a car chassis and, unlike existing cargo vans, retained the comfort and many of the other characteristics of a passenger car. Petrin estimates the value of the minivan as an innovation by implying a value to a multidimensional region of the automobile characteristic space, based on the equilibrium characteristics and prices of competing cars prior to the minivan's introduction.

As useful and important as Petrin's method is, it bears essentially no relationship to a patent race. Chrysler did not patent "the minivan," and it is highly unlikely to have been able to do so. In fact, Petrin's model presupposes that, prior to its invention, a minivan could be specified as a vector of product characteristics, whose value was determined up to some unknown taste parameter. If that were really the case, Chrysler's only innovation might have been to discover that the value of this parameter was higher than expected; in other words, the minivan was really a triumph of market research. But this is highly unlikely. It is much more likely that Chrysler's multiple patented innovations were directed to solving smaller, discrete engineering problems, such as how to modify a car's suspension to account for a minivan's higher center of gravity.[58] And because there are usually multiple ways to modify a suspension to achieve the objective, there will be multiple "owners" of any given point in the characteristic space, and therefore multiple technical substitutes. Under these circumstances, the set of innovations that is patented depends crucially on the pre-existing relationships among the competitors: if broad cross-license agreements are in effect (as is the case in the automobile industry), there is little cost to patenting an improvement that, if practiced, would infringe another firm's patent. If not, the observed set of patents will be highly selected, by inventors who are trading off the expected costs of potential liability and licensing against the expected gains from product improvement. In short, the supply of inventions is endogenous in ways that the supply of foxes is not.

Finally, even when a race to *create* patent rights can be clearly defined *ex ante*, it is rarely the case that the resulting competition defines an entire market structure *ex post*. In general, market structure depends on how patent rights, including pre-existing and subsequent rights, are *enforced*. Most economic models assume that property creation determines subsequent market structure.[59] But because the patent office does not examine patents to determine whether their practice would infringe the claims of

---

[58] In this case, the innovation lies in *maintaining* the level of one characteristic (handling) while improving the vehicle in another dimension, like hauling capacity.

[59] For example, in many economic models the winner of the first-period patent competition becomes a monopolist in second-period product competition.

any prior patent, the rights creation process cannot conclusively determine the rights enforcement process.[60] For this reason, many of the proposed reforms of the patent system (which center on improving the quality of creation, not overlaps in enforcement) would have little effect on subsequent competition in product markets.

4. Other types of intellectual property

Properly specified, the "fox" metaphor works more or less for other intellectual property systems. More importantly, the previous template offers a fairly straightforward means of comparing the economically relevant features of the patent and non-patent systems. Although non-patent systems facilitate the use and exchange of information-intensive goods, and thereby reward investment in these goods, they are less directly concerned with productivity improvement than the patent system. Therefore, a principal motivation for the comparison is to imagine alternate configurations of the patent system, using similar property models.

We can think of intellectual property generically as a 5-tuple: information; an embodiment or instantiation of that information ("subject matter"); a set of rights (to the subject matter); a jurisdiction (within which to enforce the rights); and a term (of enforcement). In the next sections, I compare the creation of three other intellectual property rights – copyright, trademark and trade secret – in these dimensions, and in other dimensions related to their exercise and enforcement. Again I focus on the dimensions that bear most closely on the investment and strategic decisions of the rights holder and his rivals. Table 1 summarizes the discussion.

4.1. Copyright

4.1.1. Subject matter

Copyright protects "original works of authorship,"[61] traditionally literary and audio/visual works; more recently, software. Insofar as works are simply consumed, copyright promotes product variety, which has some limited positive productivity impact. Insofar as the works constitute an intermediate input that reduces input costs or increases quality (a recipe, or an implementation of an algorithm in software), their productivity impact is more direct.[62] Traditionally, copyright does not protect the

---

[60] As I have previously explained, the winner of the first-period patent competition may lack the affirmative right to use his invention, so the structure of the product market depends in general on the outcome of the winner's bargaining with prior rights-holders.

[61] 17 U.S.C. §102(a).

[62] The stage of production, and the work's impact on productivity, are legally irrelevant. Like a patented invention, a copyrighted work may reduce productivity – its role is to "promote," not necessarily to effect, progress.

functional, as opposed to the informational or aesthetic, attributes of a work; copyright jurisprudence assigns functional protection to patents.[63] Often, however, there is no clear line between the form and the function of software, so "literary" works like software programs may receive substantial functional protection from copyright. The protection of software is arguably the greatest productivity-related innovation in copyright law since the dawn of the computer age.[64]

Patent law increases the public information stock by tying the disclosure of information to its embodiment in protectible subject matter. Copyright increases the public information stock by exploiting the tie that already exists between types of information. For example, the form and the content of many works are not the same, but they are typically bundled. By circumscribing the subject matter to protect one type of information but not the other, copyright induces authors to invest in the protected type, while adding the unprotected type to the public information stock.[65] Metaphorically speaking, you can patent a fox by disclosing how to obtain one; you can also copyright a fox, but not the method of obtaining one, or the idea that fox-hunting is feasible or profitable or virtuous.

4.1.2. Priority and possession

In contrast to patent law, which requires *absolute* novelty ("new"), copyright only requires *relative* novelty ("original"): two people can photograph the same building at the same moment, each creating an original, copyrightable work. Relative novelty implies that racing to establish first possession is generally unnecessary to obtain a right, however competitively advantageous. "Original" simply means "not copied."

Copyright subsists in works that are "fixed in a tangible medium of expression."[66] The definition and moment of "fixation" vary with the medium of the work and with technology,[67] but the principle is the same.

---

[63] Thus, a software implementation of a patented mathematical algorithm is also copyrighted, but it is not an infringement of the copyright to create another original software implementation of the algorithm, even if the second implementation infringes the patent.

[64] A closely related development is the digital embodiment of copyrighted works, like music and films. Digitization implies that the technology used for consuming the work (which generally requires making a copy) is closely related to the technology for reproducing the work. Much of the current debate over the proper scope of copyright law derives from the ability of a digital goods consumer to compete with the good's producer, via costless reproduction of the work.

[65] At least conceptually, the most important unprotected information type is "ideas," which are not copyrightable:

> In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.

17 U.S.C. §102(b).

[66] 17 U.S.C. §102(a).

[67] Displaying a scene on the screen of a digital camera does not fix it, but recording the scene in a computer's memory does fix it. Once the scene is fixed as an image, displaying it on the screen is a form of copying.

If "fixation" demonstrates possession of a physical thing, "expression" demonstrates possession of an eligible work of authorship. In a theatrical play, for example, protection almost always extends to the dialogue (as copyrightable expression), and almost never extends to a one-line plot summary or "pitch" (an uncopyrightable idea), but the line between abstract idea and concrete expression is hard to draw.[68]

For technology-related (and therefore productivity-related) purposes, the most important application of the idea–expression distinction is in the area of facts: facts cannot be uniquely possessed, and so are not copyrightable. The relationships among idea, expression and fact are illustrated below:

1. "Rough winds do shake the darling buds of May" (copyrightable expression).
2. "May is windy" (uncopyrightable idea).
3. "The average wind speed in May in Boston is 3.8 knots" (uncopyrightable fact).

Because facts belong to the public domain, in principle their use is non-rivalrous. Insofar as facts form the basis for productivity growth, copyright law imposes few direct constraints. On the other hand, the statute extends copyright to "compilations" – including compilations of facts – so the *access* to facts may be commercially limited.[69] Thus, the indirect productivity consequences of copyright may be significant.

The uncopyrightability of facts nicely illustrates the theme that statutory intellectual property law does not protect information *per se.* Indeed, the original justification for copyright was "to promote the progress of *science* ... by securing exclusive rights ... to *writings*...."[70] On the other hand, this means that, unlike the "useful arts" (which are stimulated by patent protection), scientific inquiry that produces facts receives comparatively little direct stimulus for its output from copyright law. For example, an early consequence of denying copyright to facts was the thin protection afforded to maps.[71] Copyright law forces a cartographer to choose between uncopyrightable

---

[68] Cf. Judge Hand's famous formulation of the common law, prior to its codification in the statute:

> Upon any work, and especially upon a play, a great number of patterns of increasing generality will fit equally well, as more and more of the incident is left out. The last may perhaps be no more than the most general statement of what the play is about, and at times might consist only of its title; but there is a point in this series of abstractions where they are no longer protected, since otherwise the playwright could prevent the use of his "ideas," to which, apart from their expression, his property is never extended. Nobody has ever been able to fix that boundary, and nobody ever can.

Nichols v. Universal Pictures Corporation, 45 F.2d 119 (2d Cir. 1930).

[69] For example, large databases typically contain copyrightable elements, such as their organization and selection of facts.

[70] U.S. constitution, Art. 1, §8, cl. 8. At the time of writing, "science" included moral philosophy and undisputed other disciplines based more on assertion (and therefore on copyrightable expression) than on statements of fact.

[71] Under U.S. law, if there is only one or a limited number of ways to express an idea, there is said to be a "merger" of the idea and its expression. Copyright law resolves this merger in favor of the public domain, by releasing the expression from its putative captivity. *Kern River Gas Transmission Co. v. Coastal Corp.*, 899 F.2d 1458, 1460 (5th Cir.) (because the idea and its expression embodied in plaintiff's maps are inseparable,

factual representation or copyrightable personal expression. Obviously, consumers demand some aspects of a map precisely to the extent that its author chooses to present facts rather than creative expression. To the extent that the author responds to that demand, he invites free-riding on his investment by those who would copy the factual, but unprotectible, elements. Similar free-riding occurs with respect to other expressions of facts that authors happen upon or invest in.

The lack of protection for pure information, and consequent underinvestment in the acquisition of that information, can also be seen in the law concerning databases and similar compilations. The initial investment in a database of facts is often substantial. Moreover, many databases derive their value by striving for comprehensive, rather than selective or otherwise creative, coverage of the facts. Thus, the U.S. Supreme Court found that an alphabetically ordered compilation of all residential telephone listings in a given region was not protectible, despite the "sweat of the brow" invested to assemble it. The Court emphasized that copyright law protects "original" works – not mechanical compilations, however costly the input or useful the output.[72] Independent creativity determines originality, which determines priority for copyright purposes.

4.1.3. Exclusivity and prior rights

Because priority merely requires independent creation, the two people who photograph the same building at the same moment can exclude others from copying their respective works, but neither can exclude the other's work, even if the works are identical. In this respect, exclusivity in copyright is fundamentally different from exclusivity in patent law, which neither permits multiple patents to be created on the same invention, nor recognizes independent creation as a defense to the enforcement of the unique right.

Like a patent, however, a copyright does not convey the right to use one's copyrighted expression: the two building photographers may be excluded from reproducing their photographs if the building itself is copyrighted, even though they own the copyrights in their photographs.

Relative to patent law, copyright law carves out a rather larger exception to the general right to exclude. In U.S. jurisprudence this exception is referred to "fair use" (17 U.S.C. §107), though similar exceptions go by other names and encompass different activities in other jurisdictions. Determining whether a given use of a copyrighted work is "fair" is a complex and essentially non-economic inquiry,[73]

"the maps at issue are not copyrightable"), cert. denied, 498 U.S. 952, 111 S.Ct. 374, 112 L.Ed.2d 336 (1990).

[72] "... the selection and arrangement of facts cannot be so mechanical or routine as to require no creativity whatsoever. The standard of originality is low, but it does exist." *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340 (1991). The Court went on to observe that "facts contained in existing works may be freely copied." *Id.* at 359.

[73] Procedurally, "fair use" is a defense to copyright infringement, not an affirmative property right. Thus, the determination of fair use follows only after a plaintiff proves that it owns a copyrighted work and that

but it does have important economic components. First, courts investigate the degree to which a second author has transformed the copyrighted work. Transformation is roughly analogous to "value added": the greater the value added by the second author, the less the market value of the derivative work is based on its use of the earlier work, therefore the more likely that the use is fair. Thus, even if the building in their photographs is copyrighted, the building photographers may be permitted to reproduce their own photographs if (say) the building is one among many buildings incidentally captured in a landscape photograph. Similarly, courts examine the degree of substitutability of the derivative work for the prior work.[74] The patent law observes a similar distinction, between infringing acts that substitute for the patent owner's marketing efforts, and infringing acts that are not substitutes. But under patent law, the finding of non-substitution does not exempt the infringer from liability; it merely reduces the penalty from the patentee's lost profits to a "reasonable royalty."

4.1.4. Notice

The general requirement for notice is increasingly attenuated in copyright law. As to the creation of rights, copyright subsists from the moment of fixation, without notice to anyone. As for enforcement of rights, affixing the copyright symbol ("©") or other notice to a work, and registering it with the Copyright Office, may afford certain benefits to the author, such as eligibility for statutory damages.

4.1.5. Institutions and jurisdiction

Because facts often have value, and because copyright law does not recognize their possession, information gatherers usually resort to non-property mechanisms of appropriating returns on their investment.[75] The availability of these mechanisms illustrates another, more general, theme: as important as the theory of property law is to a proper

---

the defendant has copied it. The inherent riskiness of asserting a fair use defense diminishes some of its economic value to users.

[74] In the United States, one of the factors that determines whether a use is fair is "the effect of the use upon the potential market for or value of the copyrighted work" (17 U.S.C. §107), which has been interpreted to mean the degree to which the derivative work substitutes for the original work. In the special case of parody, the derivative work transforms (and generally mocks) the original, sometimes to its economic detriment. Such use is generally fair; to hold otherwise would permit the parodied object to use the copyright law to stifle ridicule. Ridicule is not competition. In some instances, the copying by the derivative author may be exact and complete, but the result is not an economic substitute for the original. For example, when a second author combines the tune of Leonard Bernstein's "Maria" with politically satirical lyrics to produce "Viagra," no mockery of Bernstein or of *West Side Story* is intended. Nevertheless, "there is little or no risk of market substitution, ... and looser forms of parody may be found to be fair use...." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994), fn. 14.

[75] For example, database users almost invariably contract not to reproduce or redistribute information – including facts – that they acquire from the database, as a condition of use. The practical difference between contractual and property rights is that if A contractually obligates B not to distribute information, and B nevertheless distributes the information to C, A cannot prevent further redistribution by C, because A

understanding of intellectual property, the economic role that intellectual property plays cannot be properly understood without reference to the non-property institutions that complement, and sometimes substitute for, intellectual property.[76]

4.2. Trademark

Patent law and copyright law arise from statutes, by which the legislature seeks to balance the interests of inventors or authors with those of users. The purpose of creating these rights is to promote progress. Trademark, by contrast, is a creation of the common law. Trademark arises from the primitive common-law right for *A* to prevent *B* from marketing *B*'s goods as though they came from *A*. Such "passing off" is a species of fraud, which the common law has long prohibited. Passing off is a fraud against both the deceived consumer and the trademark owner, the latter of whom may be better able to detect it and has a greater interest in preventing it.

4.2.1. Subject matter

A trademark is a word, sign or symbol, or a combination of them, used in commerce, that refers to something else. The "something else" must include the source of a commercial good or service. For example, my computer bears "IBM" (the trademark) because IBM (the company) sold it to me. Generally speaking, anything that can designate uniquely the relationship between a good or service and its origin can serve as a trademark. For example, certain shades of pink uniquely designate fiberglass insulation (sold by Owens Corning). Owens Corning does not "own pink," nor does it even own "pink insulation." Rather, Owens Corning has established a one-to-one relationship, in the minds of consumers, between "pink" and Owens Corning the company, in the commercial field of insulation. Similarly, there is a one-to-one relationship between the character "Mickey Mouse" and the Walt Disney Company. In short, a trademark is a signal. The goal of trademark law is to preserve the one-to-one nature of the relationship (i.e., the quality of the signal), once it has been established in the minds of consumers through use in commerce.

"Use in commerce" is obviously a pregnant phrase, having a potential for ambiguity that cannot be resolved in a brief overview. The essential distinction is that, while trademarks share with company names and other indicators a signaling function, they grant control of a very limited class of information: the relationship between the seller and the thing sold. Thus, the *trademark* "McDonald's" unambiguously denotes a U.S.-based franchisor of hamburger restaurants. Whether the *name* "McDonald's" signals "fast food" or "obesity" or "American profiteering" is determined both by the ebb

---

does not "own" the information and has no contractual relationship with *C*. On the other hand, if *A* has a copyright, then as property that copyright excludes both *B* and *C* from redistribution.

[76] Other prominent examples of non-property mechanisms (and the types of intellectual property for which they substitute) are prizes (patent and copyright), research contracts (patent), non-disclosure and non-compete agreements (trade secret), and unfair competition and false advertising law (trademark).

and flow of competitive advertising, and by non-commercial use of the company and product names, over which the company has relatively little control. When assessing trademark law as a means for inducing and controlling new information, one must distinguish between a trademark proper and the much broader concept that marketing people call a "brand."

4.2.2. Location, priority and possession

On the other hand, trademarks and brands share many of the same existential attributes. The most important of these shared attributes is the controlling role of consumer perception.

Before the advent of electronic media (and, before that, high-speed transportation), consumer perception was largely determined by location. So, for example, two restaurants with the same name could co-exist commercially, without consumer confusion, if they were sufficiently separated in space. Neither could prevent the other's use of the name. The essential reason is that, if consumers could distinguish "(Abner) McDonald's of Smithtown" from "(Boris) McDonald's of Jamestown," the referent of the mark "McDonald's" was unambiguous: the relationship remained (locally) one-to-one.

The "fox" that Abner possesses is the one-to-one relationship in the minds of consumers between himself and his mark. The nature and precision of this relationship are likely to have been created by an endogenous mixture of Abner's advertising to consumers with their experience of the particular goods and services that Abner sells. Thus, the phrase "fine dining" does not distinguish Abner's restaurant from Boris's restaurant, even if it is true, while the phrase, "Our home at the bend in the river," may identify Abner's restaurant in the minds of consumers, even if "home" and "river" are highly exaggerated or entirely false. As a distinguishing phrase, the latter is eligible for protection, while the former is not.[77]

Conflict would arise if Boris wished to sell or advertise in Abner's territory. In that case, the "first in time, first in right" principle would give Abner priority in the local use of the mark. But that priority persists only as long as, and to the extent that, Abner continues to use the mark in commerce, and local consumers continue to associate the mark "McDonald's" with Abner rather than Boris. Thus, unlike a patent or a copyright, a trademark must continue in use. In other words, a trademark owner must *continue to possess* his mark. Because consumer perception fluctuates and depreciates with time

---

[77] On the other hand, if I were selling my home at the bend in a river, trademark law would deny my attempt to trademark the phrase, "Our home at the bend in the river," because the phrase is *merely descriptive* of the good offered for sale. The denial of protection to a mere description is the trademark analogue of copyright's denial of protection for a mere fact. Merely descriptive words are not eligible for protection unless the putative trademark owner can show that they additionally have acquired "secondary meaning," i.e., they are associated with a unique commercial source in the minds of consumers.

and competition, continuing possession typically requires ongoing investment, such as advertising investment – again unlike the case of patents or copyrights.[78]

Obviously, transportation and communication technologies changed this general state of affairs. Unless one plans to roll out a new product, nationwide and simultaneously, it is difficult to create a "national brand": rivals can observe the local use of the mark in commerce and use it themselves among a different population of consumers. Such imitation may not be "fraudulent" or confusing vis-à-vis any particular consumer, but it is inefficient and leads to hold-up problems as the target populations begin to overlap. The prospect of hold-up leads in turn to the familiar problem of underinvestment in new distinguishing marks.

The solution is a national system of "registered" trademarks. Like patents and copyrights, registered trademarks are created by federal statute. Much like a patent application, an application for a registered trademark claims a new word, sign or symbol. The trademark office examines the application to verify that the proposed mark will not cause confusion with other marks. Because the mark may not yet have been used in commerce, the examination is prospective, based on "objective" indicia of the mark, such as its similarity in sound or appearance to other marks in the same field of commerce. But these "objective" indicia are interpreted through the lense of (expected) consumer perception. The standard is whether there exists a "likelihood of confusion" with other marks in the same field, in the minds of consumers considering a purchase in that field. If not, the trademark office grants the application.

4.2.3. Prior rights and exclusivity

As previous sections explained, the patent office does not examine a patent application to determine if the practice of the invention it discloses would infringe any prior rights. Such examination is unnecessary: it is not an infringement to obtain a patent that, *if practiced*, would infringe a prior patent. Thus, "B infringes A" does not imply "B is unpatentable over A." This is simply another way of saying that a patent on B does not convey the right to use B.

A related point is that patent infringement is not associative: "B infringes A" does not imply "A infringes B." Similar principles apply to copyright law.

Trademark rights differ. Between potentially overlapping rights, the likelihood of confusion is typically associative: if B is likely to be confused with A, then A is likely to be confused with B. This means that the trademark office will not grant an application on B if the use of B is likely to be confused with A. Conversely, if the office grants an application for B, there must not exist a likelihood of confusion with A. In other

---

[78] Certain investments may be economically efficient but legally deficient. For example, as a matter of law, a trademarked product is one species of a larger genus: "Rollerblade" brand inline skates. It might seem to be a triumph of marketing for one firm's brand identity to dominate the competition so completely that the species serves as shorthand for the genus: "No Rollerblading." But this marketing triumph is a legal catastrophe: once consumers perceive the mark to refer to multiple sources, the mark loses the essential one-to-one relationship that is the *raison d'être* of trademark law. At that point, possession is lost.

words, unlike a patent or copyright, a trademark conveys the affirmative right to use, from inception. And it conveys symmetric rights against others, including prior registrants, to prevent their subsequent use of their earlier marks from creating a likelihood of confusion with one's own later mark.[79] This possibility leads to one final difference between trademark and other intellectual property laws: because consumer perceptions change over time, the scope of a trademark can also change, either expanding or contracting – again unlike patents or copyrights.

The exclusivity granted to a trademark owner does not extend to non-commercial uses of his mark. Here "non-commercial" is construed broadly, to allow for uses that are not deceptive or confusing, but that might cast the trademark owner in a negative light. So, for example, a trade magazine may write that, "Our tests showed that IBM's computer was 25 percent faster than Putnam's." The sentence uses two trademarks, but in a non-commercial way. IBM might then advertise its computer as, "25 percent faster than Putnam's!" – again, a permissible "non-commercial" use, because the intent and effect is to sell IBM computers (not to trade on the Putnam name), and because there is an objective, verifiable basis for the comparative statement.[80]

4.2.4. Technology

Exclusivity also does not extend to functional elements of a product, even if they are distinguishing. For example, a twist-off bottle cap might signify a one-to-one relationship with the seller of the bottle, but its functional elements are eligible for patent protection only. Similarly, a telephone number cannot be trademarked, even if unique within a commercial field, because it is also functional.

4.2.5. Notice and term

One of the reasons for restricting trademark to non-functional elements is that trademark rights last indefinitely, as long as they are not abandoned through non-use and continue to point unambiguously to a source.

Among registered marks, the examination process initially provides notice to prior trademark owners regarding potentially confusing marks. In commercial use, the superscripted designations "®" and "™" notify potential competitors of common law

---

[79] Thus, for example, the trademark office might reject "McPutnam's" as a mark in the restaurant field, as likely to be confused with "McDonald's"; on the other hand, it might accept "Putnam's." But if it does, then "McPutnam's" is probably off-limits to McDonald's as well.

[80] Unlike the determination of "fair use" in copyright law – which weighs competition between the derivative work and the original work against a finding of fair use – the law on permissible comparative advertising expressly anticipates – and defines the scope of – such competition. An economist might hypothesize that "only relative reputations matter" in consumer decision-making, but the law distinguishes between the comparative use of another's mark to promote one's own product (relative to the other), and use to disparage the other's product (relative to one's own). The latter is much more likely to be an impermissible use of the other's mark.

and registered trademarks, respectively. These notices also distinguish between trademark and non-trademark uses of identifying words ("IBM announced the release of the IBM® ThinkPad® WaitLess™ laptop. . ."), thus preserving the source-identifying function of the mark.

4.2.6. Private incentives, public policy and institutions

Much of the time, the direct productivity consequences of trademark law are probably not great. There is something to be gained from reducing consumer search costs through accurate representation of the relationship between symbols and sellers. This gain increases when product quality is not easily verified *ex ante*, or when search is inefficient.

Although reputation can act as a barrier to entry (and this barrier can be self-fulfilling when reputation depends on consumer acceptance), that is a problem with branding, not with trademarks *per se*. Similarly, the advantages of reputation can induce firms to invest in real quality improvements and other product attributes that improve reputation. Trademarks facilitate such investment, but it is unlikely that they, by themselves, induce much of it. Put differently, it is unlikely that marginal adjustments to the trademark laws would alter observed patterns of investment in brand identity.

Even when reputation is an important component of the consumer's purchase decision, finding a skilled surgeon or an honest auto mechanic is not greatly simplified by the use of trademarks. Private and public institutions – trade associations, certification boards, and government consumer protection rules – play important roles both in *ex ante* reputation and in guaranteeing performance *ex post*.[81]

The role of trademarks in the industrial organization of international trade can be quite intricate. For example, suppose that a firm sells a watch in country A, bundled with a warranty (which may be required by law), through an "authorized" distributor. Suppose that the firm sells the same watch in country B, unbundled, at a lower price. Finally, suppose that exporters in B wish to re-export the watch to A. The watch manufacturer wishes to prevent the importation of the watches, in competition with its authorized distributor (on the grounds that consumers will be confused as between the bundled and unbundled versions of the watch), or failing that, to prevent the importing distributor from implying that it is an "authorized" distributor (again on the grounds that consumers are confused as between the distributors). Depending on the jurisdiction, courts have allowed re-importation (risking the first confusion) as long as there is no misrepresentation that the seller is authorized by the manufacturer (preventing the second confusion).

---

[81] It should be noted that many private institutions that test and certify the quality of services or goods themselves rely on trademarks to distinguish their certifications and to prevent false or misleading use of the test results.

4.3. Trade secret

Trade secrets are often classified as a type of intellectual property. By treating secret information as something physically possessed by its owner, trade secrecy law directly protects information against misappropriation – an unauthorized use or disclosure in any form – by others. In that respect, it is the right most similar to traditional forms of tangible property. But because it is not primarily a means to the disclosure of information, or to its use by others,[82] it is the right least similar to other types of intellectual property.

4.3.1. Location and jurisdiction

Trade secrecy shares with trademark law an origin in the more fundamental common law of business torts. Just as one cannot appropriate information about a good's source, name or brand image for one's own use if the *good* is sold by someone else, the general rule is that one cannot appropriate information if that information: (a) was generated by another, (b) was the subject of reasonable secrecy efforts, (c) is not generally known, and (d) provides a competitive advantage.

Unlike trademark law, however, trade secret law has not coalesced around an international treaty or (in the case of the United States) even around a national standard.[83] This fragmentation complicates efforts to characterize trade secret law succinctly. To the extent that "international law" means "internationally standardized national law," trade secret law is the least "international" of all intellectual property law. Again in that respect it is most similar to the law of tangible property.

Among jurisdictions, trade secret law divides across several fault lines. First, there is disagreement over what is "generally known."[84] Similarly, there is disagreement over what it means to "appropriate" information. The law has struggled with the nonrivalrous nature of business information, which allows a second agent (such as the

---

[82] Like other types of intellectual property, trade secrecy facilitates the investment in and transfer of information. Unlike other types, it does not release information into the public domain, in exchange for limits on its use.

[83] See generally Pace (1995). An exception is the Economic Espionage Act of 1996, which provides federal criminal penalties for certain kinds of industrial spying. Although the EEA defines trade secrets broadly, and criminalizes the unauthorized copying, conveyance, receipt or possession (each broadly defined) of a trade secret (as well as attempts and conspiracies), as a criminal statute it does not provide for private actions (unlike patent, copyright or trademark infringement). For various reasons (such as the difficulty of proving criminal intent), relatively few IP-related cases are prosecuted as crimes. In 2002, 405 suspects were investigated for "IP theft," of which 92 were investigated for trade secret misappropriation (23 percent). By comparison, the number of civil cases in the United States involving patent, copyright and/or trademark infringement in 2002 was 8254. *See* U.S. Department of Justice (2004).

[84] For example, in a famous case, the Associated Press sued its competitor International News Service for re-transmitting AP's news feed to INS affiliates located in a later time zone. The U.S. Supreme Court divided over whether AP had an actual property right in the information in which it had invested (which, as facts, could not be copyrighted, and, as published locally, were not the subject of reasonable secrecy efforts), or whether INS's actions merely constituted a form of unfair competition.

spectator of a sporting event located outside a stadium) to use the information without depriving the first agent (the sponsor of the event) of his use. Since deprivation of use is the hallmark of various property crimes and torts (theft, conversion, detainer, etc.), the absence of deprivation has hindered information possessors from proving a traditional tort.[85]

4.3.2. Subject matter

Perhaps the greatest area of disagreement in trade secret law is over the information that is eligible for trade secret protection. Under traditional tort law (which governs, for example, actions in the state of New York), a trade secret "may consist of any formula, pattern, device or compilation of information which is used in one's business, and which gives him an opportunity to obtain an advantage over competitors who do not know or use it."[86] Despite the apparently broad language, this definition is very restrictive, relative to what an economist would classify as valuable private information.[87] It excludes, for example, much information that is classified as "inside information" under securities laws, on the basis of which insiders cannot trade precisely because it has value.[88] It may, in practice if not theory, exclude "negative know-how" (how *not* to do something). And it may exclude information that merely confirms what the accused misappropriator already believes, even if that confirmation reduces the risk associated with the belief.[89]

Statutory law that specifically addresses the non-rivalrous nature of trade secrets attempts to define both the information protected, and categories of misappropriation,

---

[85] Of course, appropriation deprives the information possessor of *exclusivity*, but this observation simply begs the question: is the information *property*?

[86] *Restatement of Torts*, §757.

[87] As the *Restatement* itself explains,

> [A trade secret] differs from other secret information in a business (see §759) in that it is not simply information as to single or ephemeral events in the conduct of the business as, for example, the amount or other terms of a secret bid for a contract or the salary of certain employees, or the security investments made or contemplated, or the date fixed for the announcement of a new policy or for bringing out a new model or the like. A trade secret is a process or device for continuous use in the operation of the business.

[88] Under U.S. law, "insider trading" is not limited to insiders or to shares of the insider's company.

> A company's confidential information ... qualifies as property to which the company has a right of exclusive use. The undisclosed misappropriation of such information, in violation of a fiduciary duty ... constitutes fraud akin to embezzlement – 'the fraudulent appropriation to one's own use of the money or goods entrusted to one's care by another.'

United States v. O'Hagan, 521 U.S. 642 (1997), quoting *Grin v. Shine*, 187 U.S. 181, 189 (1902) (internal citations omitted).

[89] In statistical terms, information that alters the conditional first moment of the distribution of a random variable is more easily proven to be a trade secret than information that alters higher moments.

more broadly.[90] However, even this definition does not encompass all secret "material facts."[91]

4.3.3. Exclusivity

A trade secret owner can prevent those who knew, or had reason to know, that the information is a trade secret from using or disclosing that information. If the trade secret is published, and its value thereby destroyed, the owner is entitled to damages at law. No injunction can undo the damage. But to the extent that the information remains generally secret, and is misused by another, the owner can obtain relief at equity, e.g., an injunction against further misuse or disclosure. The availability of an injunction gives a trade secret one of the key attributes of property.

In distinguishing the boundaries of exclusivity, it is important to understand the role that knowledge and intent play (in contrast to other types of intellectual property actions, which impose liability independently of knowledge or intent). Thus, someone who learns the *substance* of a trade secret must also learn (or have reason to know) its *status* as a trade secret, to be liable for misappropriation.[92]

4.3.4. Institutions

The difficulty with treating pure information as property leads most trade secret owners to protect their information contractually. A contract (a) places the recipient on notice that the information is secret, (b) explicitly circumscribes what the recipient can and cannot do with the information, and (c) expands the set of protected information to include information that may not qualify as a trade secret (e.g., employee performance evaluations). Often such agreements impose a "golden rule" condition: the recipient must treat the received information with at least the same care as it does its own information. They may also create or define a fiduciary relationship between the parties, which may impose upon the recipient "the obligation to act with loyalty and honesty and in a manner consistent with the best interests of" the owner of the confidential information.[93] Although such agreements provide important ancillary protection for

---

[90] For example, the EEA defines a trade secret as "all forms and types of financial, business, scientific, technical, economic, or engineering information ... whether tangible or intangible, and whether or how stored. . . ."

[91] A material fact is "a fact that would influence a reasonable person under the circumstances in making an investment decision (as in purchasing a security or voting for a corporate officer or action)" http://dictionary.lp.findlaw.com/.

[92] For example, suppose that A discloses information that is a trade secret to B, under appropriate confidentiality obligations. Suppose B in turn discloses to C, without informing C that the information is secret. Finally, suppose C publishes the information. A has no cause of action against C unless A can show that C ought to have known the information was a trade secret. Because information is non-rivalrous, A has difficulty monitoring B's possession of the information. For the same reason, C may have difficulty determining that the information it received from B really is a secret and that it really "belongs" to A.

[93] http://dictionary.lp.findlaw.com/.

secret information, they do not imbue the information itself with the attributes of property.

4.3.5. Title, priority, prior rights, notice and term

Because the essence of trade secrecy is secrecy, "title" and "notice" play essentially no role in trade secret law, though they may have analogues in contracts to convey the secret information. On the other hand, the temporal notions of priority, prior rights and term are not so black-and-white. Because trade secrets are simply "not generally known," there is no "first in time, first in right" principle; a second discoverer of the information has the same rights vis-à-vis third parties as does the first.[94] In permitting independent creation, both as a defense to infringement and as a source of second-party rights, trade secret law is like copyright law, but unlike patent or trademark law. But, like trademarks and unlike patents or copyrights, trade secrets may last indefinitely, depending not on use by the owner but on the absence of use (or knowledge) by others.

As a segue into the international dimension of intellectual property, it is useful to emphasize the complex interaction among the types of intellectual property and among different national rules. To take only one example: suppose that *A* discovers new information and maintains it as a trade secret, while *B* later discovers the same information and files a patent application based on it. National rules differ markedly on the respective rights of *A* and *B*, both in terms of the exclusivity granted to *B* and the conditions under which *A*'s discovery can be used as prior art against *B*'s application (see, e.g., Kupferschmid, 1993). Even within a country, the rules may differ depending on the foreign or domestic location of *A* and *B*. These relationships are highly idiosyncratic and arcane, and most of them are not derived from economic principles. But they can have enormous economic effects for unwary information investors.

5. International intellectual property

5.1. Information, coordination and "trade"

Because property rights are creatures of national governments, the international intellectual property system mostly stitches together national laws. The analytically challenging aspect of the international dimension lies not so much in the peculiarities of the international regime *per se*, but in the coordinated external effects of the national patchwork. The three most important economic features of the regime pertain

---

[94] Thus, successful reverse engineering does not necessarily destroy the trade secret. On the other hand, with each additional user the information becomes more generally known and less a source of competitive advantage. For that reason, some contracts (such as software license agreements) and some statutes (such as the Uniform Computer Information Transactions Act and the Digital Millenium Copyright Act of 1998 (P.L. 105-304)) seek to prevent most forms of reverse engineering and thereby to enhance the "priority" of the trade secret owner and the length of the trade secret's term.

to information, coordination and "trade," which I illustrate by reference to patent law but which carry over to the other types.

5.1.1. Information

Because an important purpose of intellectual property law is to add to the public information stock, many of these external effects have to do with the movement of information across national boundaries. For example, as Section 3.2 explained, the canonical justification of the patent system is the information contract between an inventor and society. I have already explained that mandating information disclosure as a condition of the "contract" is conceptually flawed. But in an international context, the notion of a "disclosure for rights" exchange is flawed for another reason: having disclosed the invention to the world in (say) his home country, the inventor provides no incremental disclosure (other than translation) by filing in foreign countries, yet his foreign applications do not fail for "lack of consideration."

National manipulation of the disclosure parameter does have international implications, however. When one country tightens its disclosure rules, thereby demanding additional "consideration," this regime change imposes externalities on all other countries. For example, in 1870 the United States began to require the inventor to disclose the "best mode" of making mechanical inventions, not just any mode; this principle was extended to all inventions in 1952 (Irving, 1992). On one hand, a German inventor benefits from the additional disclosure that U.S. patentees (not just U.S. inventors) are required to make. On the other hand, the regime change raises the cost and reduces the incentive for a German inventor to obtain a U.S. patent. If, and to the extent that, the German inventor depends on expected U.S. returns to cover the fixed cost of filing an application in his home country, the U.S. regime change may reduce the German inventor's incentive to file *in Germany*, i.e., to file at all.

Even though national borders are permeable to information, national laws do not necessarily treat domestic and foreign information the same. In the creation of patent rights, the information set against which a putative invention is evaluated depend on the information's location.[95] Similarly, the exercise of patent rights also depends on acts that occur within the jurisdiction of the entity that grants the rights, but what constitutes activity "within" a jurisdiction often varies with the jurisdiction.[96]

---

[95] For example, 35 U.S.C. §§102 (a), (b) provide that: "A person shall be entitled to a patent unless – (a) the invention was known or used by others *in this country*, or patented or described in a printed publication *in this or a foreign country*, before the invention thereof by the applicant for patent, or (b) the invention was patented or described in a printed publication *in this or a foreign country* or in public use or on sale *in this country*." Because these provisions apply equally to U.S. and non-U.S. inventors, they do not violate the "national treatment" provision of the Paris Convention (described in Section 5.2).

[96] For example, in the United States it is an infringement of a patent for a foreign firm to *offer* to sell an infringing device in the United States, even if the buyer *actually* purchases from the patentee. It is also an infringement to sell the product of a process that is patented in the United States, even if the product itself is not patented. With the advent of Internet-based business method patents, this means that a foreign Web site may infringe a U.S. patent, even though the process it employs is performed in a jurisdiction that does not recognize business method patents.

5.1.2. Coordination

The international system does impose some structure on the types of national variation that are permitted. It is easiest to understand this by way of contrast. In addition to a system in which national laws are allowed to vary freely, there are at least four types of international coordination:

1. *Homogeneity.* A patent right obtained anywhere is valid everywhere.
2. *Uniformity.* Patent rights are valid only in the country where granted, but all countries have the same standards.
3. *National treatment.* Patent rights may vary across countries, but countries may not discriminate in their treatment of domestic and foreign inventors.
4. *Reciprocity.* Countries extend patent rights to their trading partners to the same extent that their partners extend patent rights to them.

By the nineteenth century, reciprocity – effected by a system of dozens of bilateral treaties – had become the norm among intellectual property regimes. Reciprocity implies that the scope and definition of a national patent depends on the citizenship of the applicant. This feature obviously hinders the development of a consistent domestic jurisprudence: German and French inventors with U.S. patents are treated differently under U.S. law, because German law treats U.S. inventors differently than French law. The current system mixes national treatment with some limited uniformity in the creation of rights, and to a lesser extent, in their enforcement.

5.1.3. Trade

It cannot be emphasized too strongly that international intellectual property rights are not "traded," at least in the way that trade theorists think of it. There is no equilibrating price mechanism, no current account, and (given the small expenditures on patent protection, relative to national imports and exports) essentially no impact on the exchange rate. Conversely, the system creates significant departures from standard trade concepts, such that patent "trade data" are easily misinterpreted in empirical work. Section 6.2 outlines some of these departures.

5.1.4. Other types of intellectual property

There is not space to give equal treatment to the other main types of intellectual property, each of which raises its own international idiosyncrasies. But this is less of a loss than it might appear. First, only one of the three other systems – trademark – grants title by examination, and therefore generates any systematic data on the quantity of rights created. But because the value of a trademark depends, on average, on the stock of past investments in reputation, rather than on "improvement" over the "state of the art," new trademarks constitute a relatively small fraction of total value. Trademark "trade data" are therefore less informative about trademark "trade." Second, both copyright and trademark also comprise a mixture of national treatment and limited uniformity in the creation of rights, so the patent-related discussion of those

principles applies in large measure to them as well. (As for trade secret, there is no international system, so the point is moot.) Third, though many of the same information spillover issues apply in theory to copyright and trademark, they are limited by linguistic and cultural differences, which in turn limit transnational consumption and use.[97] Fourth, the role that exhaustion of rights plays in international price discrimination and in the re-importation of so-called "gray market" goods is at least as important for copyrighted and trademarked as it is for patented goods, but this issue can be analyzed much more fruitfully with traditional international trade tools. For all these reasons, I devote the bulk of the remaining discussion to the international patent system.

5.2. International institutions: patents

5.2.1. The creation of rights

The "international patent system" comprises four main treaties, administered by a variety of authorities: the Paris Convention for the Protection of Industrial Property ("Paris Convention"), administered by WIPO (1983); the Patent Cooperation Treaty ("PCT"), also administered by WIPO (1990); the European Patent Convention ("EPC"), which is administered by the European Patent Office ("EPO") (1978); and the Trade-Related Aspects of Intellectual Property Agreement ("TRIPS"), which is administered by the World Trade Organization (1995). Each of these agreements exists, in part, to standardize the patent application process, and thereby to reduce the costs of filing international patent applications. Despite the evolving standardization, each of the three later treaties actually complicates the inferences that can be drawn from the international patent micro data generated under the rules of the Paris Convention.

The PCT permits applicants to file a common application at one of a number of "international search authorities," which are located in certain national patent offices. For a relatively low fee, an applicant may designate any of the PCT member countries as a potential target. The application is examined for novelty, a relatively invariant test to determine if there exists identical prior art anywhere else in the world. If not, the application may proceed to the "national phase" in each country, which includes examination for appropriate subject matter, non-obviousness, adequacy of disclosure and other requirements that may differ by country. The primary effects of the PCT are to lower the cost of filing initial applications in a large number of countries, and to permit applicants to delay entering the national phase for up to 30 months following their priority application. Because the PCT exists in parallel with national patent systems, inventors can accomplish the same goal (a national patent) either by filing a series of individual national applications or by filing a PCT application. Obviously, the choice between these means is likely to be endogenous, depending on the value of the invention and the inventor's preference for a delayed or rapid examination. Also, because

[97] For example, most Chinese ideographs and names convey little meaning in Western countries beyond the designation (and connotations) of "Made in China."

the incremental cost of designating a country is very low, there is little information to be gained from observing such designations.

The EPC operates similarly to the PCT, in that an application designating the desired European countries is researched initially by the central EPO. However, the EPO also examines patents for all other requirements, and has the additional authority to issue patent rights that are recognized as national patents in each member country. Like the PCT, the EPC permits some applicants to economize on filing costs by paying a relatively large fixed filing fee and a relatively low per-country designation fee. The EPO is also an international search authority under the PCT, which means that applicants can file either directly with the EPO or through the PCT, designating the EPO as the search authority.

The TRIPS agreement does not directly affect the cost of filing or prosecuting a patent application. However, like the Paris Convention, TRIPS commits member countries to certain minimum standards of intellectual property (not just patent) protection. With respect to patents, the most important standardization concerns: the minimum length of protection (20 years from application date); subject matter (medicinal products must be included); and enforcement. Unlike the Paris Convention, TRIPS permits WTO members to punish the failure to enact or enforce these standards.

National patent rules generally preserve an inventor's rights from the date of application,[98] but only within that country. The Paris Convention adds to this national protection both temporal and cross-sectional safeguards for an inventor's international patent filing decision. Temporally, the Paris Convention permits applicants to wait up to one year after their initial filing date before filing an application in any other member country. The initial filing becomes the "priority filing," and the date it was filed the "priority date." Under this rule, every national application on a given invention is judged against the worldwide state of the art as it existed on the common priority date, rather than on the actual filing date in any member country.[99]

Cross-sectionally, the collection of national applications claiming priority from a single priority application is called a "patent family."[100] Loosely speaking, a patent

---

[98] By "preserving an inventor's rights" I mean that no other entity can claim priority to the invention. National rules differ as to whether another individual may use the invention before the patent application issues as a granted patent, and the extent to which a patentee may recover damages retrospectively for such pre-grant use.

[99] This grace period removes the possibility that a rival will observe the invention in one country and file applications on it ahead of the true inventor in another country (an important risk under so-called "first-to-file" regimes, which award the patent to the first filer (rather than the first inventor). Currently, the United States is the only country that awards patents to the first inventor, rather than the first filer.

[100] Given the quality of the invention, the quality of the patents on it should, in theory, be approximately constant across countries. In practice, patent quality is not the same because individual patent offices may impose different restrictions on claims. Most obviously, not all countries allow patents on certain subject matter (such as medicines or surgical methods). Even if the claims are identical, the patentee's capacity to enforce the same nominal right may also vary across target countries, depending both on target country's institutional characteristics and on transactions costs (distance, linguistic and institutional differences, etc.) that vary with the source country.

family protects a single invention in multiple countries, but national rules complicate this simple view.[101] Paris Convention rules require a policy of non-discrimination towards applicants: each country's particular patent rules (about which the Convention makes minimal stipulations) must apply equally to domestic and foreign inventors. The level procedural field at home, coupled with the additional transactions costs of filing abroad generally imply that filing at home is cheaper.

Uniformity standardizes the definition of an application, patent examination, and patent enforcement, across countries. The PCT is a uniform application regime, while the EPC is a uniform examination regime. Although the concept of a European or world patent has long been discussed, there exist no international enforcement standards (such as the definition of infringement), nor is there a competent court having international jurisdiction.

Because filing at home is cheaper, and because no cost or risk arises from waiting up to a year to file abroad, most applicants file their priority application in their home country, then wait until the one-year anniversary of their initial filing to decide in which foreign countries they should also file. In theory, the country composition of a patent family is completely determined on the one-year anniversary of the priority filing.[102]

Although TRIPS has caused patent standards to converge somewhat, the patent screening process differs markedly across countries. In most countries, all patent applications are published, independent of the patent office's decision to accept or reject them. In many countries, the patent office publishes an application again if and when it is granted, to reflect any changes (such as modifying the claims or augmenting the disclosure) that the office may have required. In some countries, notably Germany and Japan, applicant must pay a separate fee to initiate formal examination proceedings; he has up to seven years from filing to exercise this option. Many countries also require that the applicant pay increasing annual renewal fees; otherwise the application (or, if granted, the patent) lapses permanently. After the national patent office allows the patent, some countries permit other interested parties to oppose that decision, before the application is actually granted (Graham *et al.*, 2003). Thus, the screening process

---

[101] In any given country, a patent application may result in zero, one or multiple issued patents for a given invention, depending respectively on whether the patent office rejects, allows or divides the application.

[102] The inventor may still file equivalent applications in foreign countries, while not claiming priority from his initial application. Such applications are not considered part of the patent family. Moreover, such applications are examined in light of the additional prior art that has accumulated since the priority application, which lessens the likelihood that they will be granted in their entirety. Once at least one country has published the inventor's application (typically 18 months after filing), it becomes prior art against which all subsequent applications worldwide are judged. If one country has already published an inventor's application, other countries will treat that publication as prior art and automatically refuse to grant a patent on equivalent applications. Thus, an inventor who fails to file abroad within the one-year grace period risks substantial diminution of his foreign patent rights; after 18 months these rights are almost certainly lost.

reflects endogenous choices by the applicant, and perhaps his rivals, in addition to the exogenous patentability standards imposed by the national patent office.[103]

5.2.2. The enforcement of rights

Nearly all of the international standardization has occurred in the areas of creation and examination. Once granted, patent rights are subject to national precedent and procedure. Even under the European Patent Convention, post-issuance enforcement (including patent infringement litigation) remains under the jurisdiction of the individual member countries. In short, the *de jure* convergence of rights is limited by their *de facto* interpretation.

National control over patent enforcement does more than reinforce the regime's patchwork nature. It also has direct, but uncertain, consequences for trade in goods and services. In most countries, a patentee who sells the patented product "exhausts" his rights with the initial sale, which means that the buyer can resell in competition with the patentee without infringing the patent. The exhaustion doctrine limits price discrimination within a country. Across countries, however, the issues are less clear-cut (Stack, 1998). Suppose that an inventor in source country S obtains patent protection in targets T and U and exports the patented product to each country, selling at different prices. Suppose that the price in T is lower, and that the T importer (or his customer) decides to re-export to consumers in U (sometimes called "parallel importation" into U). Such attempted arbitrage would violate the inventor's exclusive rights to sell and to import into U, *if* the inventor's initial sale to T does not exhaust his rights in U. Consumers in U may advocate an expansive interpretation of the exhaustion doctrine under U's laws, in the belief that competitive imports from both S and T reduce the price they must pay. On the other hand, if T allows re-export and U allows parallel imports, S's optimal response to such arbitrage is likely to be higher prices in T, or even a corner solution: the inventor may refuse to sell to T and may then sell to U at pre-arbitrage prices (depending on demand and cost parameters).[104] The end result is that U's interpretation of the exhaustion doctrine partially determines whether and at what price the patented product is sold in T.

---

[103] The most important of these exogenous standards are: (1) protectible subject matter (the patent laws protect "inventions," not "discoveries"); (2) novelty (regardless of field, the invention cannot previously have been disclosed in a prior document or offered for sale); (3) non-obviousness (within the relevant field, the invention must not be obvious to a person of ordinary skill applying routine procedures); (4) utility (the invention must be capable of industrial application); (5) enabling disclosure (the application must enable others of ordinary skill in the relevant field to make the invention (including, in some countries, the "best mode" of the invention contemplated by the inventor as of the filing date)). In the United States, *see* 35 U.S.C. §§101, 102, 103, and 112.

[104] These are not simply theoretical concerns. Developed countries, especially the United States, are under severe pressure from domestic consumers to permit parallel imports of otherwise expensive pharmaceuticals, such as HIV/AIDS treatments. U.S. pharmaceutical companies have, in turn, threatened not to sell to certain target markets (most notably Canada) that permit re-exportation. That threat has prompted Canadian counter-threats of compulsory licensing, which would effectively remove from the patent holder the right to exclude others.

The WTO regime explicitly permits each importing country to decide whether or not it will permit parallel imports.[105] But because some of the costs of each country's decision may fall on other countries (in the form of higher prices or zero quantity), the current regime is unlikely to remain stable.

6. International trade and intellectual property

At the outset, I explained that property rights are primitives that exist prior to economic activity. It should not be surprising, therefore, that such economic content as is found in the process of creating intellectual property rights must generally be inferred indirectly, from "quantity" data. The purpose of this section is to show how some of those inferences can be drawn properly in the contexts of national income accounting and of international trade.

National patent offices do not grant patents based on their economic value, but on their technical merit.[106] Overall, this distinction is useful, because it keeps subjective claims of private value from contaminating the objective legal/technical determination of patentability. The lack of correlation between technical merit and private value complicates both the measurement of private value and the formulation of public policy.[107] It also complicates national income accounting.

6.1. National income accounting

An applicant must only satisfy minimum technical standards to obtain a patent. This administrative threshold is a lower bound on the quality of the invention. Similarly,

---

105

> For the purposes of dispute settlement under this Agreement, ... nothing in this Agreement shall be used to address the issue of the exhaustion of intellectual property rights.

TRIPS, Article 6 (*Exhaustion*).

106 Over time, U.S. (and other) courts have come to accept certain "objective [economic] indicia" as indirect evidence of the technical merit of an invention (beginning with *Graham v. John Deere*, 383 U.S. 1 (1966)). Some of these, such as "long-felt need" and "commercial success," are explicitly demand-related; others, such as "evidence of the failure of others" and "acquiescence" by rivals (taking licenses under the patent) depend on supply-side decisions. These indicia generally are not introduced during the application process, but only in litigation, to rebut an accused infringer's defense that the patent is obvious.

107 There are varied reasons for the imperfect correlation between private value and technical merit. (1) A patent's claims may not cover embodiments or uses of the invention, because (a) the claims may have been drafted based on imperfect information about future configurations of the invention; (b) rivals possess private information about other embodiments and uses; (c) the patent's disclosure leads to "creative destruction" of the benefits of its claims, as imitators discover how to design around them (Schumpeter, 1942; Caballero and Jaffe, 1993); (2) patent rights are specific assets, for which markets are thin and transactions costs are high, and so have value that depends in general on their owner (Teece, 1986); (3) inventions often demand complementary assets to generate commercial success: a technically superior product, divorced from a feasible manufacturing process, has little private value.

in an optimizing framework the application must separately satisfy a minimum economic threshold. Because the application's technical merit and economic value are not perfectly correlated, and because the observations on quality and value are censored at their respective thresholds, the input cost of obtaining patent protection is a noisy and downwardly biased measure of the patent's market value.

From a national income perspective, censoring means that the resources invested in obtaining patents systematically understate the value of national patent portfolio. The "value added" – the difference between the value of inputs into patenting and the value of the output – is retained by the input owners. If one conceived of a national "patent market," the excess of domestic production value over domestic production cost could be treated as inframarginal rents earned by patentees. But, unlike a competitive market, the administrative process does not permit other input suppliers, particularly the government patent office, to charge prices that reflect their value added or the scarcity of the inputs. For that reason, the aggregate value of national patent rights is unobserved. And in any event, it is more accurate to think of the "supply" of patent applications instead as the demand for a public service, derived from and conditional on realizations of (largely) private R&D processes.[108]

The possibility of international trade further complicates the accounting issues. First, the measurement of patent trade is complicated by the definition of the home country. In some countries, such as Canada, applicants often file abroad first (in the United States), then wait until the one-year anniversary to decide whether to file in their home country. As Putnam (1996) shows, this choice is endogenous. It also need not reflect the source of the R&D resources or competence, which may (for example) have originated in the corporate parent of the foreign subsidiary that filed the application.

But even assuming that the home country is identified correctly, other "identification" issues arise. For example, it is common to refer to the patents obtained in target country T by inventors from source country S as "exports" from S to T. From the point of view of national income accounting, the export of a patent right is really the purchase of services from a foreign government, i.e., it is an import.[109] This imported service creates an asset, held by S, which may be thought of as a contingent claim. But it is not necessarily a claim by S against T. For example, suppose the S firm sues a firm from U for infringement of its patent in T. Then S's claim is against U, enforceable in the courts of T. S's patent in T constitutes a claim against T only to the extent that the patent raises prices to T's consumers and/or reduces the profits of T's producers. The multilateral nature of S's claim in T, and the selectivity employed in choosing which of S's patents to file in T, together vitiate the notion of a "technology

---

[108] It should be clear that R&D expenditure decisions are not exogenous, or even necessarily prior, to patent application decisions. Most of the "D" part of R&D occurs after the patent application has been filed. Hall *et al.* (1986) find that firm-level patent counts peak approximately contemporaneously with R&D expenditures, and the lag structure of the relationship is relatively flat.

[109] Despite this conceptual misstatement, the usage is so deeply ingrained that I continue to follow it below.

supply curve" from S to T. Formally, S inventors demand services from T to enforce claims against U.

It should be noted that most of the subtleties involved in accounting for international patent rights are attributable to the introduction of (notional) patent "prices" (i.e., patent values), overlaid on top of the observed patent quantities.[110] In the absence of prices, it seems more analogous (or at least more traditional) to associate the purchase of a patent right abroad with the (potential) export of the underlying technology, and I follow that tradition here: (1) "domestic production" is the quantity of domestic patents generated by each country's residents; (2) "imports" refers to the quantity of domestic applications the country receives from non-residents; and (3) "exports" refers to the quantity of equivalent foreign applications on the inventions that constitute "domestic production." But this terminology can be misleading, so it is helpful to establish some initial notation.[111]

6.2. The international patent data generating process

In what follows, I refer generically to "an invention." As a matter of law, an invention does not exist unless a patent application is granted. Under a purely legal definition, therefore, the same technological advance is considered "an invention" in some countries but not others. To minimize both ambiguity and sample selection, I define an invention to be "*any* candidate technological advance that has been the subject of at least one patent application anywhere in the world." Contrary to the legal definition of "invention," this economic definition implies that "an inventor" and "an applicant" are the same thing. "A patentee," on the other hand, is the recipient of a granted patent in a particular country. I define an invention's source country to be the country where the first patent application is filed.

6.2.1. Notation

For the ith invention generated in source country s, $s = 1, \ldots, S$, let $\mathcal{P}_{ist}$ be a random variable, the realization of which determines the value of filing for patent protection in target country t, $t = 1, \ldots, T$. If the inventor files for protection, then $1[\mathcal{P}_{ist}] = 1$,

---

[110] In addition to their impact on the multilateral issues raised above, the introduction of prices can reverse the sign of the "current account." In Table 2, the United States is the largest "exporter" of patents (measured by the quantity of patents that U.S. inventors obtain abroad). But Putnam (1996) finds that the *value* of the foreign rights held by U.S. inventors is less than the value of the rights that the U.S. government sells to foreign inventors. In that sense, the United States runs a "patent value deficit" (i.e., it is a "net importer" of patent rights), mainly because the larger U.S. economy implies that the right to exclude others from that economy has greater value.

[111] Apart from the subtleties introduced by prices, there are several basic differences between the standard international trade framework and the international patent filing process. For example: the same patent application can be (1) filed at home and abroad ("domestic production" and "exporting" are not mutually exclusive), (2) exported to more than one country, and/or (3) granted in some countries but not others.

where $1[\cdot]$ is an indicator function. Denote by $\mathcal{P}_{st} = \sum_i 1[\mathcal{P}_{ist}]$ the number of applications filed by inventors from s in t.[112]

It will be helpful to distinguish domestic applications on domestic inventions ($\mathcal{P}_{ss}$) from others. I define $\mathcal{R}_s \equiv \mathcal{P}_{ss}$ as "resident applications," or "resident inventions." (Under the definition I have employed, the relationship is one-to-one.)

Let $\mathcal{E}_s = \sum_t \mathcal{P}_{st}, t \neq s$ denote exports from country s to all other target countries t ("export applications"). Export applications $\mathcal{E}_s$ count separately the applications on a single invention in multiple foreign countries. Thus, national patent application production $\mathcal{Y}_s$ is equal to the sum of resident applications and export applications:

$$\mathcal{Y}_s = \mathcal{R}_s + \mathcal{E}_s. \tag{1}$$

On the import side, let applications to target country t from all source countries be $\mathcal{M}_t = \sum_s \mathcal{P}_{st}, s \neq t$. Where no ambiguity will ensue, I refer to country s's exports as $\mathcal{E}_s$ and its imports as $\mathcal{M}_s$.

6.2.2. *"Macro" issues*

Now, consider the construction of a patent "current account." For country s the national consumption of patent applications $\mathcal{C}_s$ is the sum of resident applications and import applications:

$$\mathcal{C}_s = \mathcal{R}_s + \mathcal{M}_s. \tag{2}$$

Therefore, combining (1) and (2), national "patent savings" is

$$\mathcal{Y}_s - \mathcal{C}_s = (\mathcal{R}_s + \mathcal{E}_s) - (\mathcal{R}_s + \mathcal{M}_s) = \mathcal{E}_s - \mathcal{M}_s. \tag{3}$$

In other words, national savings is equal to the current account deficit. I refer to a current account measured in *applications* as a "type 1 deficit." A type 1 deficit may be positive or negative.

For many productivity-related purposes, this "deficit" is misleading. On the consumption side, each resident application ($\mathcal{R}_s$) and each imported application (in $\mathcal{M}_s$) represent a unique invention; the relationship between applications and inventions is one-to-one. But on the production side, patent exports $\mathcal{E}_s$ represent multiple equivalent foreign applications generated by each resident invention; the export relationship is many-to-one. For that reason, Equation (3) subtracts imported (unique) *inventions* $\mathcal{M}_s$ from exported (non-unique) *applications* $\mathcal{E}_s$. From a productivity perspective, this

---

[112] Putnam (1996) follows the "deterministic" patent renewal literature (e.g., Schankerman and Pakes, 1986) by defining $1[\cdot]$ as $V_{0it} \gtrless C_{0it}$ (where V_{0it} is the present value of annual returns to patent protection in t net of the annual maintenance fees, depreciated at a fixed (estimated) rate until the year in which the maintenance fee exceeds the annual return, at which point the patent lapses; and C_{0it} is the initial cost of filing in t (which, in general, depends on distance and language and so varies with s)). With extensions to the multi-country context, $1[\cdot]$ could be adapted to follow the "stochastic" renewal literature (Pakes, 1986; Lanjouw, 1998), e.g., $E[V_{0it}|r_{1it}] \gtrless C_{0it}$, where $r_{1it}|\alpha_i$ is the initial annual return in t (conditional on the common (across target countries) invention "quality" draw α_i), which evolves stochastically.

is an apples-and-oranges comparison: there is no reason to think that importing two inventions is balanced by exporting the same invention twice.[113]

A more intuitive way for economists to think about net national savings and international trade is to compare national *invention* production $\mathcal{R}_s$ with national *invention* consumption $\mathcal{C}_s$. But then net national savings is

$$\mathcal{R}_s - \mathcal{C}_s = \mathcal{R}_s - (\mathcal{R}_s + \mathcal{M}_s) = -\mathcal{M}_s < 0. \tag{4}$$

In other words, every country must consume more inventions than it produces.[114] I refer to net national savings measured in inventions as a "type 2 deficit."

An obvious trade-based objection to Equation (4) is that it does not separately identify inventions intended solely for the domestic market from those that are intended for export. But this objection simply highlights the differences between trade in goods and trade in patents. Denote by $\mathcal{R}_s^-$ inventions that are filed only at home ("domestic-only inventions") and by $\mathcal{R}_s^+$ inventions that are filed at home but also filed in at least one foreign country ("export inventions").[115] By definition, $\mathcal{R}_s \equiv \mathcal{R}_s^- + \mathcal{R}_s^+$. Because "domestic production" and "exports" are not mutually exclusive patent activities, so-called export inventions $\mathcal{R}_s^+$ may also be used at home. As a corollary, there is no such thing as a pure export, because the inventions in $\mathcal{R}_s^+$ are (by definition) protected first at home – even if the expected value of filing in the home market is negative.[116]

Given an estimate of the number of (unique) export inventions $\mathcal{R}_s^+$, we could compute a true current account (or "type 3 deficit"),

$$\mathcal{R}_s^+ - \mathcal{M}_s \gtrless 0, \tag{5}$$

but $\mathcal{R}_s^+$ cannot be observed in the macro trade data.

In short, despite the natural inclination to analogize a patent trade deficit to a goods and services deficit, macro-data based calculations derived from either application counts (3) or invention counts (4) have no logically consistent interpretation, while the "true" current account (5) is unobservable. For this reason, the use of macroeconomic indicators may lead to badly mistaken inferences about the flow of patent value into and out of the country.

---

[113] For example, McCalman (2005) employs data on patent imports (inventions) to explain patent exports (applications).

[114] It is important to recognize that this deficit arises independently of, and in addition to, the inventions that a country may import without compensation because the foreign inventor did not file for patent protection in the importing country.

[115] Here a "+" superscript denotes patenting activity abroad by resident inventors.

[116] Under the option model of Putnam (1996), an export-oriented producer nevertheless may file first at home, despite the negative expected value of doing so, because the one-year delay permitted under the Paris Convention permits the inventor to file at home (where filing is also likely to be cheaper) and to gather additional information about his invention's profitability before filing abroad, without risk to the priority of the foreign applications. Using that model, it is possible to simulate $\mathcal{R}_s^+$ and to remove from $\mathcal{R}_s$ those patent applications having negative expected value in the home country. Of course, once one has simulated the *value* of domestic and foreign patent rights, the adjustment to patent *quantities* is often superfluous.

6.2.3. Data

We can illustrate many of these macroeconomic and international trade issues by examining an old cohort of patent applications. Table 2, columns (1)–(5), show $\mathcal{P}_{st}$, the average annual "trade" between source and target countries among the top-five patenting countries, averaged over the years 1972–1976 (WIPO, 1983, 1990).[117] These cohorts represent the relatively simple international patent regime before the PCT and EPC dramatically increased the number of countries in which inventors file, as well as the complexity of their filing calculus, by introducing parallel regional and global application routes to patent protection.[118]

The data show, for example, that U.S. inventors filed an average of 65,293 resident applications ($\mathcal{R}_{\text{US}}$) each year. These applications resulted in an average of 45,031 applications in the four other top-five countries ($\mathcal{E}_{\text{US}}$), for a top-five export ratio of 0.69. Again, the export ratio represents multiple exports of the same inventions. By comparison, the United States imported an average of 24,767 inventions ($\mathcal{M}_{\text{US}}$).[119]

The top-five countries imported approximately the same total number of inventions from each other (21,000–26,000). But the import total masks important differences in the sources of imports. These differences show up more clearly in the total export statistics: the United States exported almost 60 percent more applications than

---

[117] The top-five countries were the United States, the former West Germany, Japan, the United Kingdom, and France. These countries were the most frequent export targets (accounting for 44 percent of world imports) and the most prolific exporters (accounting for 73 percent of exports). The next leading countries (Italy and Canada) averaged about six percent of imports each, and about two percent of exports. Note that Canada is the United States' largest patent trading partner.

[118] The existence of multiple patent application avenues, and multiple examination standards, considerably complicates the modeling and interpretation of European patent data. (Putnam (1996) analyzes the effect on optimizing behavior of overlaying an EPO-style regional patent application regime onto national regimes.) Expanding inventor filing options also has sharply diminished the quality of the innovation signal generated by international patent data. For example, over the 1972–1976 period, Belgium received an average of 14,475 applications annually, and granted an average of 14,404 patents. (Under its patent registration system, Belgium granted essentially 100 percent of the applications it received.) About 7.5 percent of these were granted to Belgian residents.

By 2000, the data show the number of Belgian applications increasing nearly ten-fold, to 141,766, of which about 37.5 percent designated Belgium through the EPO, and the remainder through the PCT. However, the number of patents *granted* in Belgium in 2000 actually fell 16 percent from the 1972–1976 average, to 12,122 (an 8.6 percent grant rate). Belgian residents received about 6 percent of all granted patents, for a total decline in resident patents of 36 percent. Thus, "real" inventive activity (actual patent grants) appears to have declined moderately over this period, particularly among Belgian inventors. Because of the EPO's more stringent examination standards, and its option to delay examination, inventors face a higher quality threshold to obtain a patent, and they face it with better information about the true quality of their invention. Thus, it is not clear whether an actual decline in resident inventive activity has occurred. But it is clear that "nominal" application counts bear little relationship to "real" inventive activity.

[119] Because $\mathcal{M}_{\text{US}}$ are unique inventions and $\mathcal{E}_{\text{US}}$ are not, the type 1 deficit ($\mathcal{E}_{\text{US}} - \mathcal{M}_{\text{US}} = +20{,}264$) is non-sensical. The data do not disclose the number of resident *inventions* that were exported, $\mathcal{R}^{+}_{\text{US}}$; from the macro data themselves, we can only infer that $\mathcal{R}^{+}_{\text{US}} \geqslant \max_{t}[\mathcal{P}_{\text{US},t}] = \mathcal{P}_{\text{US,CANADA}} \approx 15{,}500$.

Table 2. Average annual patent applications, by source and target country ($\mathcal{P}_{st}$), and related export statistics. Top-five countries, 1972–1976

	Export data					Export statistics					
	Target country t $\mathcal{P}_{st}$ $\mathcal{P}_{ss} \equiv \mathcal{R}_s$					Exports $\mathcal{E}_s$	Export ratio μ_s	Export probability π_s^+	Export inventions $\mathcal{R}_s^+$	Type 3 deficit	Conditional export ratio μ_s^+
Source country s	U.S.	Germany	Japan	UK	France	$= \sum_t \mathcal{P}_{st,t \text{ not } s}$	$= \mathcal{E}/\mathcal{R}_s$	(est.)	$= \pi_s^+ * \mathcal{R}_s$	$= \mathcal{R}_s^+ - \mathcal{M}_s$	$= \mathcal{E}_s/\mathcal{R}_s^+$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
U.S.	65,293	11,105	12,203	12,410	9312	45,031	0.69	0.46	30,020	5254	1.50
Germany	8377	31,417	5250	6979	7741	28,347	0.90	0.61	19, 017	(2826)	1.49
Japan	8498	4629	121,788	3882	2652	19,661	0.16	0.07	8653	(12,757)	2.27
UK	4786	3122	2309	21,999	2680	12,897	0.59	0.38	8287	(17,682)	1.56
France	3106	2987	1649	2698	12,910	10,440	0.81	0.56	7195	(15,191)	1.45
Top 5 total $\mathcal{M} = \sum_s \mathcal{P}_{st,t \text{ not } s}$	24,767	21,843	21,410	25,969	22,386	116,375	0.46	0.29	73,173	(43,202)	1.59

	Conditional trade probabilities $= \pi_{st}^+/\mathcal{R}_s^+$					All top-five probability	
						π_5: complete independence	π_5: complete dependence
Source country s	(12)	(13)	(14)	(15)	(16)	(17)	(18)
U.S.	–	0.37	0.41	0.41	0.31	0.02	0.31
Germany	0.44	–	0.28	0.37	0.41	0.02	0.28
Japan	0.98	0.53	–	0.45	0.31	0.07	0.31
U.K.	0.58	0.38	0.28	–	0.32	0.02	0.28
France	0.43	0.42	0.23	0.37	–	0.02	0.23
Top 5 total	0.57	0.40	0.33	0.40	0.34	0.02	0.29

Sources: Data: *Industrial Property Statistics* (WIPO), 1972–1976. Statistics: Putnam (2007).

the next-largest exporter (Germany), and almost 4.5 times as many applications as France.

Because of differences in both national patent rules and the size of the domestic economy, one cannot readily compare patenting levels or export rates from one source country to the next. For example, at the time of the sample Japan permitted only one claim per patent, whereas the other countries allowed multiple claims. Thus, Japan generated over 121,000 domestic applications (nearly twice the U.S. level), but its inventors patented in the three European countries at less than half the level of U.S. inventors. These figures imply, for example, that the probability that a U.S. inventor filed in the United Kingdom was about 0.19 (= 12,410/65,293), but the comparable estimate for a Japanese inventor was only about 0.03.

Similarly, the combination of national idiosyncracies and the lack of structural modeling frustrates efforts to create simple patent indicators. For example, the top-five export ratios ($\mu_s \equiv \frac{\mathcal{E}_s}{\mathcal{R}_s}$), reported in column (7) of Table 2, varied from 0.16 (Japan) to 0.90 (Germany). But it is impossible to tell whether German inventors are really more "productive" than Japanese inventors, or whether the differences simply are due to national patent rules (which determine $\mathcal{R}_s$) or to some other factor unrelated to inventive productivity.

6.2.4. "Micro" issues

These macroeconomic accounting issues raise analogous microeconomic issues when trying to describe and interpret optimal patenting behavior. For example, one explanation for the variation in μ_s is (unobserved) invention quality: countries that create more valuable inventions are likely to export them more widely. But another explanation could be sample selection: the export ratio depends on the share of inventions patented abroad. If country s has a small domestic market, then resident inventors will file for protection at home only if there is a high probability that they will also file abroad. Alternatively, if inventors from country s face higher transaction costs abroad (due, say, to distance or language barriers), then the minimum returns threshold required to justify filing abroad will be higher. This implies that the probability of filing abroad may be lower, but the mean patent quality (conditional on filing abroad) may be higher.

To make these hypotheses precise, it is helpful to add some further notation. Recalling that total resident inventions $\mathcal{R}_s$ is the sum of domestic-only inventions ($\mathcal{R}_s^-$) and export inventions ($\mathcal{R}_s^+$), we can define $\pi_s^+ \equiv \mathcal{R}_s^+/\mathcal{R}_s$ to be the proportion of resident inventions that are exported. Similarly, let $\mu_s^+ \equiv \mathcal{E}_s/\mathcal{R}_s^+$ be the average number of countries in which an inventor from country s files an application, conditional on export. I call π_s^+ the "export probability" and μ_s^+ the "conditional export ratio."

We can then create an export identity as follows:

$$\mathcal{E}_s = \frac{\mathcal{E}_s}{\mathcal{R}_s^+}\frac{\mathcal{R}_s^+}{\mathcal{R}_s}\mathcal{R}_s \tag{6}$$

or

$$\mathcal{E}_s = \mu_s^+ \pi_s^+ \mathcal{R}_s. \tag{7}$$

Equation (7) decomposes the national export ratio μ_s into the product of two effects, unobserved heterogeneity in invention quality (as reflected in the conditional export ratio μ_s^+) and sample selection (as reflected in the export probability π_s^+):

$$\mu_s \equiv \mu_s^+ \pi_s^+. \tag{8}$$

Because $\mathcal{R}_s^+$ cannot be observed in the macro data, neither μ_s^+ nor π_s^+ can be computed directly. Discriminating between unobserved heterogeneity and sample selection either requires micro data (invention-level data that reveal whether an individual invention is exported or not), or an estimate of either μ_s^+ or π_s^+.

6.2.5. *Explaining international patenting behavior*

Putnam (2007) develops econometric techniques for estimating π_s^+ using only macro data. The result is shown in column (8) of Table 2. For example, Japan is estimated to export only about 7 percent of its inventions, while Germany and France export over half of theirs. Thus, the sample of internationally patented inventions is highly selected, and the selection rules appear to differ markedly across countries.

Given π_s^+, one can estimate the number of export inventions by $\mathcal{R}_s^+ = \pi_s^+ \mathcal{R}_s$; the calculation is shown in column (9). For example, the export probability for the United States is estimated to be about 0.46, which implies that $\mathcal{R}_{\text{US}}^+ = (0.46)(65{,}293) = 30{,}020$. As column (9) shows, ranking countries by $\mathcal{R}^+$ provides a much clearer picture of their international inventive capacity than does ranking by resident applications $\mathcal{R}$.

After identifying $\mathcal{R}_s^+$, one can construct a more natural invention-based "current account." The type 3 deficit for each country is shown in column (10). Among the top-five countries, only the United States showed a "trade surplus" in inventions; the remaining countries produced "deficits" ranging from an average of about 10 percent of domestic production (Germany and Japan) to an average of about 100 percent (the United Kingdom and France).[120]

Since $\mu_s^+ = \mu_s / \pi_s^+$ (or equivalently, $\mu_s^+ = \mathcal{E}_s / \mathcal{R}_s^+$), one can readily compute the conditional export ratio from columns (6)–(9); the result is shown in column (11). Again Japan is the outlier: conditional on export, Japan averages about 50 percent more top-five exports (2.27) than do the remaining countries (≈ 1.50).[121] Apart from Japan, μ^+ falls in a remarkably narrow range around its mean. This result shows

---

[120] Even under this definition, there is no requirement of globally "balanced trade." For example, in the aggregate, the top-five countries show an annual deficit of almost 43,200 inventions (= 73,200–116,400), or $\mu_5^+ = 1.59$ imports for every export. The reason, again, is that each country exports the same invention multiple times.

[121] Putnam (2007) demonstrates that the high estimated value of μ_{JP}^+ is not simply driven by the low estimate for π_{JP}^+. Micro data on individual inventions show that Japanese inventors are almost twice as likely to patent in all of the top-five countries as would be expected from their share of world export inventions ($\mathcal{R}_{\text{JP}}^+ / \mathcal{R}_{\text{W}}^+$), while they are less than half as likely to patent in any 10 or more countries, given their export share. In other words, much more than do their main rivals, Japanese inventors concentrate their foreign patenting in the other four top-five countries.

that most of the variation in the naive export ratio μ (column (7)) appears to be due to sample selection (i.e., variation in π^+) rather than to unobserved heterogeneity.[122]

Another important advantage of identifying $\mathcal{R}_s^+$ is that one can properly compute pairwise conditional export probabilities $\hat{\pi}_{st}^+ \equiv \mathcal{P}_{st}/\mathcal{R}_s^+$. For example, given the estimates in column (9), the implied probability that a U.S. inventor files in the United Kingdom, conditional on export, is $12{,}410/30{,}020 = 0.41$.

The complete 5×5 table of export probabilities is shown in columns (12)–(16) of Table 2. The columns of this table show remarkable consistency in exporting behavior. For example, the export probabilities for Germany and France – two neighboring countries that do not share a common language with any of their top-five trading partners – are very close to being the same for each of the other three countries. But the main advantage of the table is that it removes the bias introduced by Japan's unique domestic patent rules, and highlights Japan's behavior towards the other top-five countries: Japanese inventors have the highest probability of filing in each of its partners (except France), conditional on export, among the top-five countries. In the United States, that probability approaches one.

Aggregated over all importing countries, the weighted (by export inventions) conditional probabilities correct the (mistaken) inference from the common level of imports in the raw data that each country is approximately equally attractive as an export target. For example, the United Kingdom imported about 1200 more inventions than the United States, but the conditional probability of filing in the United States (0.57) is nearly 50 percent higher than that for the United Kingdom (0.40).

6.2.6. The foundations of a structural model

Because these macro data comprise quantities but not prices, they remain in the realm of "patent indicators" (Griliches, 1990). Rather than rely on "indicators," accurate inference requires a structural model of international patenting behavior, based on prices. Such a model generates direct valuations of international trade in intellectual property rights, and provides further corrections to naive, indicators-based inferences,[123] but it lies beyond the scope of a "law and economics" introduction.

For present purposes, I again emphasize the distinction between exporting intellectual property rights and exporting goods: unlike goods, export and domestic consumption are not mutually exclusive activities; again unlike goods, exports to target

---

[122] Putnam (1996) develops a structural model of foreign filing, which explains whether inventors file abroad and where they file when they do, and which underlies the observed export probability. The delay permitted by the Paris Convention introduces an option value to filing at home; the price of this option is the price of a domestic filing, net of the expected domestic returns. The effect of this option value is to render domestic filing profitable even when domestic filing has negative expected value, and (in certain circumstances) even when foreign filing also has negative expected value (at the time of domestic filing). What matters to the foreign filing decision is the rate at which the inventor learns about the invention over time, i.e., whether the foreign filing option is "in the money" at the one-year deadline. Since domestic filing costs partly determine the cost of the option, and since the domestic patent office helps determine the

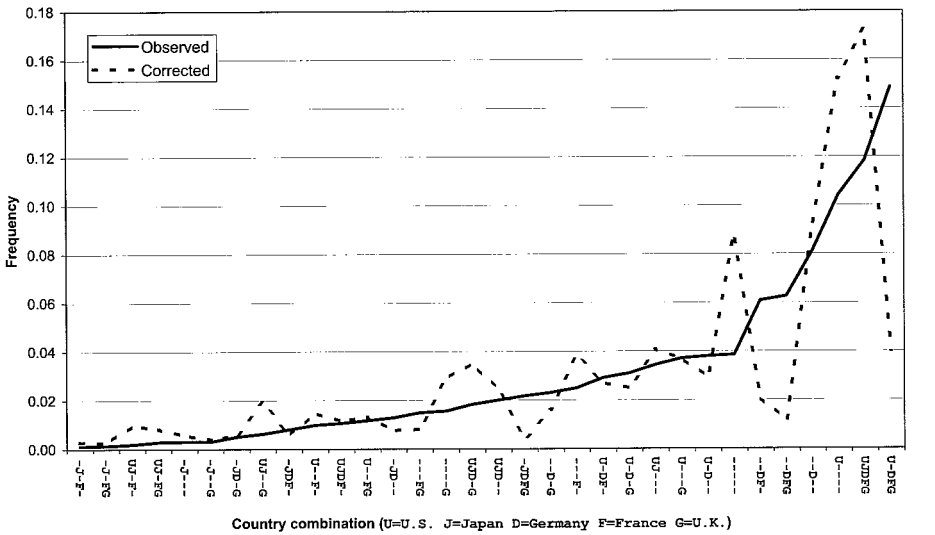

Fig. 1. Observed and corrected application frequencies for combinations of the top-five countries, 1974 cohort.

countries t and u are not mutually exclusive. For these reasons, optimal patent export behavior is characterized by a *combination* $c \in C$ of countries in which to obtain intellectual property rights, where C is the set of all 2^T possible target country combinations. The traditional pairwise representation of patent flows (columns (1)–(5) of Table 2) masks this behavior.

Figure 1 summarizes some of the additional detail on country combinations, which can only be obtained from patent micro data. I have plotted the empirical frequency (conditional on export) π_c^+ of each of the $\|C\| = 2^5$ possible combinations of the top-five countries (including "zero" combinations that are patented entirely outside the top-five). The frequencies are ordered from least to greatest, from the 1974 world cohort. The data show a moderately diffuse distribution: for example, inventions filed in all top-five countries ("UJDFG") account for only about 12 percent of the sample; I denote the estimated probability of this combination by π_5^+.[124]

---

inventor's rate of learning and thereby the value of the option, domestic patent rules strongly influence the observed probability of filing abroad.

[123] For example, Table 2 implies that the United States runs a substantial "trade surplus" with the rest of the world, as measured by the *number* of inventions exported and imported. But, using the micro data analogue of Table 2, Putnam (1996) shows that the United States grants greater *value* in patent rights to foreigners than the value of the rights it purchases abroad – i.e., it runs a "trade deficit."

[124] Note that the most frequent international filing combination in the micro data is U-DFG (i.e., the United States plus the three European countries), which accounts for about 15 percent of all combinations observed.

To set this value in context, it is helpful to compare it with the implications of the macro data. Suppose that returns to patent protection are completely independent across target countries. Then, for inventors from source country s, the probability of filing in all $t \neq s$ is $\prod_{t \neq s} \hat{\pi}_{st}$. This calculation is shown in column (17) of Table 2. Despite the moderately different target country probabilities for each source country, the estimates of π_5^+ are close to 0.02 for all source countries except Japan (0.07).

Alternatively, suppose that returns to patent protection are completely dependent across target countries. In that case, differences in the observed probabilities arise solely from differences in the cost of filing and target market size. Then the estimate of π_5^+ is $\min_{t \neq s}\{\hat{\pi}_{st}\}$. These calculations are shown in column (18) of Table 2. Again, the estimates of π_5^+ are close to 0.30 for all countries except France (0.23).

Although each model generates similar results for each source country, each set of estimates diverges from the observed average value for π_5^+ (0.12). The divergence implies that returns to patent protection (and therefore the filing probabilities $\hat{\pi}_{st}$) must be neither completely independent nor completely dependent across target countries.

Putnam (1996) develops a structural model that incorporates unobserved heterogeneity (the common invention quality across target countries), sample selection (the probability of entering the sample of export inventions) and idiosyncratic target country characteristics as determinants of the behavior observed in the micro data. Putnam (2007) discusses other sample selection mechanisms (most importantly, the U.S. practice of publishing only granted applications) and other reporting anomalies (particularly affecting Japanese data) which imply that the frequencies observed in the micro data are biased.

Using those insights, Putnam (2007) develops techniques with which to estimate combination probabilities using only the macro data reported in Table 2, by optimally combining calculations of the type shown in columns (17) and (18) of Table 2. The corrected results are plotted in Figure 1. In general, Figure 1 shows that combinations including Japan are underrepresented in the micro data. Most notably, corrections based on the macro data increase the estimate of π_5^+ by 45 percent, from the observed value of 0.12 to 0.17, making the combination UJDFG the most frequent observed. These corrected combination probabilities then enter the maximum likelihood estimation of the structural model (whose unit of observation is the combination of target countries for each source country) as optimal observation weights.

6.2.7. Spillovers

Lastly, a neglected implication of Table 2 and Figure 1 is how *little* international patenting actually occurs. The diagonal of Table 2 shows that the top-five countries produced $\mathcal{R}_5 \approx 253{,}400$ inventions per year during this period. Had they all been protected in all five countries, these inventions would have generated $\bar{\mathcal{E}}_5 = 4 \times 253{,}400 = 1{,}013{,}600$ export applications. In reality, the sum of column (6) shows that $\mathcal{E}_5 \approx 116{,}400$ applications, or about 0.11 [$= \mu_5/4$] of the potential maximum. Even if one restricts the sample to the $\mathcal{R}_5^+ \approx 73{,}200$ export inventions (the sum of

column (9)), which could have produced up to $\bar{\mathcal{E}}_5^+ \approx 292{,}800$ export applications, the result is $\mathcal{E}_5/\bar{\mathcal{E}}_5^+ [= \mu_5^+/4] < 0.40$ of the potential maximum. As the sample of countries is expanded to include less frequent targets, these ratios decline markedly: over 24 leading countries, Putnam (2007) estimates the unconditional and conditional ratios at about 0.04 and 0.13 of the potential maxima, respectively. Since, in the aggregate, only about half of all applications are granted, this means that the world's inventions led to granted patents in fewer than 0.02 (i.e., two percent) of all possible cases.

Given the usual policy justifications for the patent system, it is natural to identify these holes in patent protection with spillover opportunities. While there is no doubt that spillovers occur, the data highlight an important ambiguity in their interpretation.[125] To the extent that the "person of ordinary skill in the art" – whose skill level probably varies, as a matter of law, from country to country – is actually incapable of reproducing the invention, even after reading the patent's disclosure, type (1) spillover is legally, but not economically, feasible.[126] If an inventor assigns low value to patent protection in a given target country because he already effectively enjoys exclusivity there (i.e., he need not fear competition from potential imitators for whom the disclosure requirement is too weak to actually enable imitation), then the patent system in that country has failed to create *either* private value (from additional exclusivity) *or* public value (from additional disclosure).

6.3. Other types of intellectual property data

As I have explained previously, the empirical analysis of other types of intellectual property is hampered by the absence of systematic data. The international trademark regime collects data on new national filings, which are published in WIPO's *Industrial Property Statistics* serial. But as yet there has been no effort, as there has been with patents, to develop an optimizing model of trademark acquisition, renewal or "trade."

---

[125] Spillover occurs with or without the filing of an application in any particular country. But the nature of the spillover is different, at least in theory. If no application is filed in a given target country, spillovers should shift the production function (a "type (1)" spillover), by enabling legally permissible copying of the disclosed *claims*. Spillovers also shift the R&D search function of rival firms (a "type (2)" spillover), whether or not an application is filed in any particular target country, by shifting the demand and lowering the costs of further search built on the disclosed *information*. Many empirical analyses do not distinguish carefully between these two types of spillovers.

The low rate of international filing implies that, conditional on the unavoidable type (2) spillovers, the value of preventing type (1) spillovers in additional countries is small, because (a) most inventions are not worth copying directly, (b) most inventions cannot be copied directly even after disclosure, and/or (c) the incremental benefits of protection, *conditional* on protection in other trading partners, are low.

[126] Within the TRIPS framework, it would be permissible for a developing nation to tighten disclosure requirements to reflect the lower skill of its resident inventors, as long as these requirements were applied equally to resident and foreign inventors. But, as in the discussion of the "best mode" requirement under U.S. law (Section 5), tightening the disclosure requirement would have obvious external effects that would substantially increase the costs of filing in that country.

Moreover, any model of trademark depreciation (which depends primarily on *own* advertising investments, and which usually results in sustained *appreciation*), must depart fundamentally from that of patent depreciation (which depends primarily on *competing* technological investments, and which usually results in obsolescence). It must also account for the potentially (and endogenously) infinite life of a trademark, rather than the exogenously finite maximum patent life. Finally, many of the world's most well-known and highly appreciated marks pre-date the collection of trademark filing data.

Copyrights and trade secrets are subject to opposite empirical problems. With copyrighted works, everything from a student's classroom notes to the next *Mona Lisa* is protected from the moment of fixation. Since every work is protected, there is no identifying event and therefore no "sample" generated. Moreover, many works, such as computer software, are updated regularly, which means that the definition of what their copyright covers is continuously evolving. Under these circumstances, the aggregate value of copyrights is inferred – very imprecisely, and probably with considerable bias – by the owners of "copyright-intensive" goods, like books and films (e.g., WIPO, 2003, and national studies cited therein). Needless to say, associating the value of an input with the output that embodies it begs the question of how much value the output adds to the input.

Conversely, trade secret valuation suffers from the obvious empirical difficulty that trade secret protection depends on non-publication, and the less obvious difficulty that trade secrets are rarely quantified or even defined. Empirical economists might make some progress using (out-of-date) micro data on firm decision-making with respect to internal employee "invention disclosures," on some of which the firm pursues patent protection, while it maintains others as trade secrets, and rejects still others. But there appear to be no large-sample efforts along these lines thus far.

7. *Conclusion*

Most economists instinctively seek productivity growth. That search extends to finding more efficient ways to implement systems that "promote the progress of science and the useful arts, by securing ... exclusive rights...." This paper strongly supports that objective. Yet, insofar as they derive from the law of property, intellectual property and related systems share a centuries-old legal and empirical framework. This framework has exhibited remarkable adaptability and resiliency to a wide variety of resource types, including information. At least in the near term, efficiency improvements are likely to come by further adaptation of the principles of property. At the same time, the optimal implementable innovation system might well not be a "property" system at all, but some other system engineered from the first principles of law and economics (Wright, 1983).

Even if first principles can produce the optimal system, it is unlikely that the world trading system will gravitate to that system without extensive empirical analysis. As the brief summary of Section 6.2 indicates, the non-rivalrous nature of intellectual

property "exports" complicates the interpretation of international intellectual property rights "trade" data. Having in hand a structural model adapted to measuring the *value* – not just the *quantity* – of intellectual property rights, economists can employ traditional trade-based empirical techniques to determine which national characteristics – and in particular, which national legal rules – translate that value into the greatest impact on productivity growth.

Whether better systems (and, especially, *globally* better systems) are adapted from property law or cut from whole cloth, it is important for economists to identify their legal parameters. Like Morse's telegraph patent, it is not enough to *claim* that such systems exist; economists must *describe* their implementation. Those parameters, and that implementation, can best be evaluated in a law-and-economics perspective by their answer to a simple question: "who" "gets" "the fox"?

References

Caballero, R., Jaffe, A. (1993), How high are the giant's shoulders? An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In: Blanchard, O., Fischer, S. (Eds.), *NBER Macroeconomics Annual, vol. 8*. MIT Press, Cambridge.

Chisum, D. (serial), *Chisum on Patents*. LexisNexis Matthew Bender, New York.

Deardorff, A.V. (1992), Welfare effects of global patent protection. *Economica* 59, 35–51.

Dinwoodie, G., Hennessy, W., Perlmutter, S. (2001), *International Intellectual Property Law and Policy*. Lexis Publishing, New York.

Dinwoodie, G., Hennessy, W., Perlmutter, S. (2003), *International and Comparative Patent Law*. Matthew Bender, New York.

Dwyer, J., Menell, P.S. (2001), *Property Law and Policy*. Foundation Press, New York.

Eaton, J., Kortum, S. (1996), Trade in ideas: Productivity and patenting in the OECD. *Journal of International Economics* 40, 251–278.

Eaton, J., Kortum, S. (1999), International technology diffusion: Theory and measurement. *International Economic Review* 40 (3), 537–570.

Gallini, N., Scotchmer, S. (2002), Intellectual property: When is it the best incentive system? In: Jaffe, A., Lerner, J., Stern, S. (Eds.), *Innovation Policy and the Economy, vol. 2*. MIT Press, Cambridge.

Gilson, J., LaLonde, A. (serial), *Gilson on Trademarks*. LexisNexis Matthew Bender, New York.

Goldstein, P. (2001), *International Intellectual Property Law: Cases and Materials*. Foundation Press, New York.

Graham, S.J.H., Hall, B.H., Harhoff, D., Mowery, D.C. (2003), Post-issue patent 'quality control': A comparative study of US patent re-examinations and European patent oppositions. *Working Paper No. 8807*. National Bureau of Economic Research, Cambridge.

Griliches, Z. (1990), Patent statistics as economic indicators: A survey. *Journal of Economic Literature* 28, 1661–1707.

Hall, B.H., Griliches, Z., Hausman, J.A. (1986), Patents and R&D: Is there a lag? *International Economic Review* 27 (2), 265–284.

Hovenkamp, H., Janis, M.D., Lemley, M.A. (2002), *IP and Antitrust: An Analysis of Antitrust Principles Applied to Intellectual Property Law*. Aspen Law and Business Publishers, Frederick, MD.

Irving, T. (1992), The significant federal circuit cases interpreting section 112. *The American University Law Review* 41 (3).

Jackson, J. (2002), *The World Trading System: Law and Policy of International Economic Relations*, second ed. MIT Press, Cambridge, MA.

Kupferschmid, K. (1993), Prior user rights: The inventors lottery ticket. *AIPLA Quarterly Journal* 21, 213–228.

Lanjouw, J.O. (1998), Patent protection in the shadow of infringement: Simulation estimations of patent value. *The Review of Economic Studies* 65 (4), 671–710.

Lanjouw, J., Pakes, A., Putnam, J. (1998), How to count patents and value intellectual property: The uses of patent renewal and application data. *Journal of Industrial Economics* 46 (4), 405–432.

Lemley, M.A. (2001), Rational ignorance at the patent office. *Northwestern University Law Review* 95 (4), 1495–1532.

Lemley, M.A., Shapiro, C. (2005), Probabilistic patents. *Journal of Economic Perspectives* 19 (2), 7598.

Lerner, J. (2002), 150 years of patent protection. *American Economic Review* 92 (2), 221–225; see also the working paper version. Available at: http://www.nber.org/papers/w8977.

Lerner, J. (2005), 150 years of patent office practice. *American Law and Economics Review* 7 (1), 112–143.

Masoudi, G. (2006), Promoting economic development through sound competition policy. *Taiwan 2006 International Conference on Competition Policies/Laws, Taipei, Taiwan, June 20, 2006*. Available at: http://www.usdoj.gov/atr/public/speeches/217616.htm. (Visited June 19, 2007.)

McCalman, P. (2005), Who enjoys 'TRIPS' abroad? An empirical analysis of intellectual property rights in the Uruguay Round. *Canadian Journal of Economics* 38 (2), 574–603.

Merges, R., Duffy, J. (2002), *Patent Law and Policy: Cases and Materials*, third ed. LexisNexis Matthew Bender, New York.

Merges, R., Menell, P.S., Lemley, M.A. (2006), *Intellectual Property in the Technological Age*. Aspen Law and Business Publishers, Frederick, MD.

Milgrim, R. (serial), *Milgrim on Trade Secrets*. LexisNexis Matthew Bender, New York.

Nimmer, D. (serial), *Nimmer on Copyright*. LexisNexis Matthew Bender, New York.

Pace, C.R.J. (1995), The case for a federal trade secrets act. *Harvard Journal of Law and Technology* 8 (2), 427–469.

Pakes, A. (1986), Patents as options: Some estimates of the value of holding European patent stocks. *Econometrica* 54 (4), 755–784.

Petrin, A. (2002), Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy* 110 (4), 705–729.

Putnam, J. (1996), The value of international patent rights. *Unpublished Ph.D. dissertation*. Yale University.

Putnam, J. (2007), International patenting. *Unpublished manuscript*. University of Toronto, Institute for Policy Analysis.

Putnam, J., Tepperman, A. (2004), Revisiting the cost of bad patents: For whom is rational ignorance rational? *Intellectual Property Today* (October), 37–39.

Putnam, J., Tepperman, A. (2005), Do Intellectual Property Rights promote economic progress? A review of the empirical literature. *Unpublished manuscript*. CRA International.

Reinganum, J. (1981), Dynamic games of innovation. *Journal of Economic Theory* 25, 2141.

Schankerman, M., Pakes, A. (1986), Estimates of the value of patent rights in European countries during the post-1950 period. *Economic Journal* 96 (384), 1052–1076.

Schumpeter, J. (1942), *Capitalism, Socialism and Democracy*. Harper & Row, New York.

Scotchmer, S. (1991), Standing on the shoulders of giants: Cumulative research and the patent law. *Journal of Economic Perspectives* 5 (1), 29–41.

Shapiro, C. (2001), Navigating the patent thicket: Cross licenses, patent pools, and standard-setting. In: Jaffe, A., Lerner, J., Stern, S. (Eds.), *Innovation Policy and the Economy, vol. 1*. MIT Press, Cambridge.

Stack, A. (1998), Trips, patent exhaustion and parallel imports. *The Journal of World Intellectual Property* 1 (4), 657–689.

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

Trebilcock, M.J., Howse, R. (2005), *The Regulation of International Trade*, third ed. Routledge Press, Oxford.

U.S. Department of Justice (2004), Intellectual Property Theft, 2002. Federal Justice Statistics Program. Available at: http://www.ojp.usdoj.gov/bjs/abstract/ipt02.htm. (Visited June 15, 2007.)

U.S. Department of Justice and Federal Trade Commission (1995), *Antitrust Guidelines for the Licensing of Intellectual Property*.

U.S. Department of Justice and Federal Trade Commission (2007), *Antitrust Enforcement and Intellectual Property Rights: Promoting Innovation and Competition*.

Weigmann, S. (2000), Researching Intellectual Property Law in an International Context. Boston University School of Law. Available at: http://www.llrx.com/features/iplaw.htm. (Visited June 15, 2007.)

World Intellectual Property Organization (1983), *Industrial Property Statistics*. Geneva.

World Intellectual Property Organization (1990), *Industrial Property Statistics*. Geneva.

World Intellectual Property Organization (2003), *Guide on surveying the economic contribution of the copyright-based industries*. Geneva.

Wright, B. (1983), The economics of investment incentives: Patents, prizes, and research contracts. *American Economic Review* 73, 691–707.