An Efficient Breach Theory of the *Numerus Clausus*

Contact Information:
Leah Theriault
Washington University School of Law
Campus Box 1120
One Brookings Drive
St. Louis, MO 63130
314-935-7857
ltheriault@law.wustl.edu

Abstract

Limitations on the customizability of property rights (the *numerus clausus* principle) are a puzzling feature of the common law conception of property. An economic rationale, drawing on the notion of efficient breach, is offered to explain the modern persistence of the doctrine. Application of the *numerus clausus* principle limits the proliferation of governance rules in the economy, replacing them with rules of exclusion. Rules of exclusion allow possessors of assets to unilaterally seek out new uses of those assets. Governance rules either specify uses, or require possessors to remit payment or get permission for new uses. Whenever the law replaces a governance rule with a rule of exclusion, efficient breach has the potential to occur because possessors of assets can apply their unique, rival and nontransferable human capital inputs to tangible assets, generating outputs (the new uses) that move resources to their higher-value uses. This is how all innovation, both high-tech and low-tech, occurs. Governance rules generate lower levels of innovation because they separate, to a greater or lesser degree, possession from use rights. This separation introduces the problem of asymmetries in the cognition and motivation of the possessor of the asset and the owner of the use rights. Unlike information asymmetries, cognitive and motivation asymmetries are not cured by notice, and can hinder the search for the higher-value uses of resources. This theory's focus on search, innovation and human capital explains why the *numerus clausus* principle remains most robust in the areas of personal and intellectual property, and why it has been somewhat attenuated in the areas of real property and financial instruments.

Introduction

"[I]t must not … be supposed that incidents of a novel kind can be devised and attached to property at the fancy or caprice of the owner."[1]

[1] Keppell v. Bailey 39 Eng. Rep. 1042 (Ch. 1834).

Why are contract rights, in the eyes of the law, so much more customizable than property rights? Why, once we allocate a property right, should we not give that private party complete freedom to determine its use and disposition forever – even after it has passed into the hands of another? Freedom to divide it up into as many strands as possible, to transfer it into the hands of another with whatever use restrictions are desired, to allow a free range of uses by transferees but demand payment for each and every use? Freedom of contract would seem to support such a state of affairs, and the general economic view sees no good reason to impose such a limitation on a private right. Indeed, economics generally supports the imposition of servitudes, for example, to encourage property owners to make the investments necessary for the efficient development and use of resources.[2] Still others point to the fact that present owners have a greater moral claim on enforcing their preferences vis-à-vis their property than future, often unknown, owners. Property law, on the other hand, interferes with the private delineation of property rights almost as a matter of course, and places limits on the number of property forms which can be created by private parties.

This judicial interference with the will of the owner creates a '*numerus clausus*' - which literally means 'closed number' – a limited menu of property forms from which private parties must choose when they seek to structure their economic activities. This term may sound alien to those versed in the American common law (and perhaps it should, since the term is borrowed from civil law systems), but the principle that it represents is ubiquitous throughout property law. Property scholars will recognize it in the guise of the principle against restraints on alienation (these restraints create novel, highly restrictive forms of property), and the infamous maze of servitude requirements (which limit the ability to create new kinds of servitudes). Intellectual property scholars will be most familiar with it in its 'license or sale' incarnation: the recurring debate about how robust the doctrine of first sale should be in the face of technological or organizational innovations that alter the costs of notice. But the principle runs throughout all of property law, and is particularly pivotal in policing the line between contract and property rights. Indeed, it is one of the underlying themes of this work that the issue of the scope of property rights, the issue of servitudes,[3] and the 'license or sale' question, are at heart one issue, which is better addressed by a general theory - under the *numerus clausus* umbrella – than by isolated justifications or treatments.

Does all this legal interference with the will of private actors make good economic sense? The *numerus clausus*, which is so central to the universe of property law,[4] has, until recently, been largely ignored by both economists and 'law and economics' scholars.[5] There are at present two principle economic theories[6] of the *numerus clausus*: 1) the third party information costs argument developed by Thomas Merrill and Henry Smith; and 2) the costs of verifying divided rights rationale proffered by Henry Hansmann and Reinier Kraakman. These theories are important and seminal contributions to the literature, and they have significantly advanced our understanding of the *numerus clausus*, but they do not account for two persistent features of the law. First, they focus on the impact of novel forms of property on third parties rather than on those within the chain of privity. The law, in contrast, is squarely concerned with the impact on those directly involved with the novel property form. Second, both theories explicitly state that as technology lessens the cost of providing

[2] See, e.g., Carol Rose, Servitudes, Security and Assent
[3] Which has a huge literature behind it. See, e.g., Parisi, Sterk, French, etc.
[4] Even though it is largely implicit in the common law (unlike the civil law). See Part 1.
[5] See, e.g., Merrill & Smith.
[6] There is actually a third, the anticommons theory put forward by Michael Heller. This theory is subject to the challenge that the law actually allows significant fragmentation of property interests. For this reason, it will not be discussed here. Although the efficient breach theory is, like Heller's, concerned with underuse, in the efficient breach theory underuse does not arise from excessive fragmentation, but from any separation of possession from use rights. For a brief review of the various theories, see Heller, The Economy of Property Forms. Also see Heller, Rudden, and Hansmann & Kraakman. There are an increasing number of non-economic theories as well, but these will not be dealt with in this work.

notice of novel forms, the law's restrictions on the number of forms should also decrease: "In general, to the extent that technological change allows cheaper notice of relevant interests, the need for standardization by the law will be somewhat diminished".[7] But the law, while allowing a certain amount of relaxation of the *numerus clausus* upon the advent of effective recordation systems, continues to enforce the principle even in situations where notice is cheap and effective. To address these lacunae, this work constructs a justification of the *numerus clausus* in which the principle 1) focuses on the impact that novel forms of property have on persons within the chain of privity; and 2) does not vary inversely in its robustness with advances in technology (or organization) that reduce the costs of notice.

One thing is certain: technological advances and organizational innovations will continue to alter the ways in which novel property forms can be created and enforced. We need only look at the capability of digitized databases to provide relatively quick and cheap notice of unusual use restrictions on real estate, for example, to see the enormous potential for reduction in information and notice costs. And intellectual property scholars are all too familiar with the ways in which Digital Rights Management (DRM) systems can completely overwhelm the legally constructed boundaries of both copyright and contract law. This happens because technological restrictions on digitized content actually become part of the product that is transferred to the consumer,[8] and because digital goods can provide notice of contractual restrictions on content each time the product is accessed. Given that technology continues to reduce not only the cost of providing notice of novel restrictions, but also the cost of monitoring and enforcing those restrictions, any theory of the *numerus clausus* that turns on third party information costs will be subject to ever more severe erosion over time. The erosion that began in the real property area with the development of land registries continues apace in the digital arena. Digitized goods can provide notice of novel restrictions by presenting a 'click-thru' set of terms every time the good is used, because the content cannot usually perceived without running some kind of computerized device. The software within this device can easily be coded to display terms of use, and to require assent to them, as a condition of accessing the content or running the program. Those who take refuge in the safeguards that contract has traditionally provided to defeat any legally undesirable restrictions should consider that judges are ever more willing to recognize new forms of assent by action in order to bring contract law into the digital age: "Formality is not a requisite; any sign, symbol or action, or even willful inaction, as long as it is unequivocally referable to the promise, may create a contract."[9] Organizational innovations, such as licensing collectives and institutional licensing, similarly allow for the creation of incremental payment schemes which are often the defining feature of restrictive forms of property. The time is ripe, therefore, to ask whether the *numerus clausus* can be economically justified in the face of the decreasing transaction costs that inevitably accompany advances in technology and organization. Can the principle be truly, and completely, freed from the tyranny of notice?[10]

Furthermore, in the case of digital 'things', the legal tools formerly used to force the transfer of an unrestricted property bundle do not even have a chance to operate.[11] Indeed, the digitally-encoded restrictions themselves become 'part of the product', and absent some sort of legislation that

---

[7] Merrill & Smith at 42. References removed.
[8] See, e.g., Robinson, Van Houweling, etc.
[9] Those interested in judicial eloquence might want to read the preamble to the quote: "Promises become binding when there is a meeting of the minds and consideration is exchanged. So it was at King's Bench in common law England; so it was under the common law in the American colonies; so it was through more than two centuries of jurisprudence in this country; and so it is today. Assent may be registered by a signature, a handshake, or a click of a computer mouse transmitted across the invisible ether of the Internet."
[10] The strongest advocate of the view that effective notice should allow unilateral creation of all manner of novel property forms is Richard Epstein.
[11] Radin et al.

mandates product design,[12] it is increasingly difficult for contract law to address these new forms of restrictive property. In fact, recognition that this class of goods represents a significant upsurge in the prevalence of servitudes on chattels has led some commentators to call for a re-examination of the traditional legal stance of disallowing them.[13] If third party information costs are really the driving force behind the law's antipathy to new property forms, then notice that operates on the face of the good, that even goes so far as to present itself every time a good is used, removes the need for the doctrine completely. This form of notice reduces information costs to those both within and without the chain of title, because the need to search for unusual restrictions in any particular case is eliminated as the use of these automated tools becomes ever more common. And although these tools are currently more likely to be used in the context of digital content or electronic goods (i.e., any device which uses a computer chip can either display terms or electronically control at least some uses), innovations in remote sensing devices, such as Radio Frequency Identification tags (RFIDs) and nanotechnology, present the possibility that even conventional goods might some day be subject to electronic servitudes. Indeed, there is some evidence that producers of tangible products are already trying to use license terms to impose restrictions on the use of their products.[14] Though such methods are low-tech, they do provide evidence of a desire on the part of manufacturers to control their products post-sale. The possibility that use restrictions and repeated payment systems could become the norm, rather than the exception, is therefore genuine. And so, it is time to ask whether the law's traditional interference with such 'fancies' is a mere relic of days past, when notice of novel restrictions (and their monitoring and enforcement) was simply too expensive for sellers to incorporate into their transactions. This lack of technological capability had the effect of creating both a *de facto* and *de jure* transactional form which permeates, and literally defines, the conventional and legal understanding of 'property'. But now that this technological hurdle has been breached, should our traditional, perhaps unsophisticated, legal constructs also be cast away? In other words, is 'sale'[15] nothing more than a technological artifact?

In the interests of clarity, and because the topic of these pages crosses so many contentious doctrinal and theoretical fault lines, a few words need to be said up front on what this work does and does *not* assert. First, it does not assert, as some others have, that the common law is necessarily efficient. What this work does assert is that a *structural background of exclusionary rules* is efficient. The fact that the law has traditionally fostered this structural background is a separate matter, albeit one that makes the legal rule economically justifiable. Second, and related to the first point, this work makes no statement as to relative institutional competence. Instead, it focuses on the *type of rules* that promote efficiency. This departs from the predominant approach, which revolves around an assessment of the costs and benefits of institutional choice. A substantial amount of legal and economic scholarship is concerned with assessing whether governments, markets, courts or firms are better at gathering and processing the information necessary to make efficient decisions. In contrast, the perspective developed here is completely agnostic on the question of institutional choice.

This work is comprised of five parts: Parts 1 and 2 develop the theory; Parts 3 and 4 respond to the principle objections to the theory that arise from the existing literature; and Part 5 applies the theory to selected policy issues. Part 1, *'What is the Numerus Clausus? The Law's Preference for Rules of Exclusion'*, gives an overview of the current economic theories of the *numerus clausus*, and offers an alternative[16] conception which equates the principle with a legal preference for rules of

[12] Much in the same way that legislation regulated the 'product' of insurance contracts when egregious terms became the industry norm.
[13] Robinson, Personal Property Servitudes
[14] Actually, such attempts have a venerable history behind them. E.g., Edison.
[15] And the *numerus clausus*, and the legal preference for exclusionary rules.
[16] Note that all of these definitions are complementary, rather than antagonistic. See Part 1.

exclusion, rather than governance rules. Exclusionary rules allow possessors of assets to unilaterally seek out new uses of those assets. Governance rules either specify uses, or require possessors to remit payment or get permission for new uses. By forcing full and free use of an asset to travel with the asset when it permanently changes hands, the *numerus clausus* operates to replace governance rules with rules of exclusion. This allows the transferee of an asset to freely breach governance rules imposed on an asset by the transferor. Thus we see that breach has a place in property law, as well as contract law.

Part 2, *'User Innovation – When Breaching Governance Rules is Efficient'*, discusses when the replacement of governance rules with rules of exclusion results in a form of *efficient* breach. Replacing a governance rule with a rule of exclusion will be efficient whenever users of assets search for new, higher-value uses of those assets. These searches will be highly variable across individuals because each person's stock of human capital, which is a critical input into the search for new uses of all resources, is rival, non-transferable and unique. To optimize the process of constrained maximization, therefore, we need a device that manages the inevitable trade-off between providing sufficient incentives to present property owners (past possessors), and allowing future owners (present possessors) to apply their unique inputs in the search for new uses. The *numerus clausus* provides such a device: by replacing governance rules with rules of exclusion whenever a tangible asset permanently changes hands, transferees of assets become owners of assets, free to unilaterally determine the uses of those assets without compensation.

Part 3, *'The Third Party Issue - From the Price Mechanism to Proxies,'* addresses the first major objection to the efficient breach theory that arises from the existing literature. In contrast to the two extant theories, which focus on the information costs to third parties, the justification put forward in this work focuses on the parties within the chain of title, or 'zone of privity'.[17] The standard objection to this approach - that the costs to parties within the chain of title are accounted for by the price mechanism - is dealt with head on. Property (a rule of exclusion) is nothing other than the supersession of the price mechanism, with the controlling mind of the firm owner replaced by the controlling mind of the property owner. The *numerus clausus*, by imposing a rule of exclusion on the parties, thus operates to supersede the price mechanism. The question we need to answer is therefore: *why* does the *numerus clausus* supersede the price mechanism? Having disposed of the primary objection of the standard account, this section then moves on to assess the relation between the supersession of the price mechanism and the pivotal role played by tangibility in the law. It concludes that the law uses *tangibility as a very rough proxy for human capital*. Unlike most proxies, this proxy is more effective precisely because it is rough. The reason for this is simple. It is impossible to identify *ex ante* whether any one decision-making unit will have, or will employ, the human capital that will produce these new uses. Property law, via the *numerus clausus*, therefore assumes that there will be a new use every time an asset permanently changes hands.

Part 4, *'Notice Does Not Cure Asymmetries in Cognition and Motivation'*, addresses the second major objection to the new theory: that effective notice should result in a relaxation of the *numerus clausus* principle. The answer to this objection is found in the fact that the *numerus clausus* seeks to compensate for asymmetries in cognition and motivation, rather than asymmetries in information. In contrast to information asymmetries, which are completely resolved by notice, asymmetries in cognition and motivation persist even after all parties are fully informed of the parameters of the new property form. There is no reason, therefore, for the law's restriction on the creation of new property forms to turn on the cost of providing notice of those forms. This new theory thus accounts for the fact that the *numerus clausus* persists in the law even where notice is cheap and effective Although it is impossible to identify the 'perfect' mix of governance and exclusion in the economy (i.e., exactly how 'closed' or 'open' the *numerus clausus* should be), rules

[17] The general nature of the principle does produce system-wide benefits, however.

of exclusion are indicated whenever there will be benefits to searching out new uses of assets. It is to be expected, therefore, that the *numerus clausus* will be (and should be) more persistent in personal and intellectual property, and somewhat attenuated in real property and financial instruments.

Part 5, *'Applying the Theory'*, discusses some theoretical implications of the efficient breach theory of the *numerus clausus* for contract law and economics. It then applies the theory to selected legal policy issues. The legal policy applications include several ongoing controversies in the fields of personal and intellectual property. The following are discussed in some detail: disputes about Digital Rights Management (DRM) in intellectual property; the first sale concept in copyright and patent law; new licensing methods that seek to extend post-sale control over tangible products; and recent scholarship questioning the policy against servitudes on chattels.

The Conclusion briefly reviews the argument, and sets out a preliminary agenda for future research. Of particular importance is the question of when exclusionary rules are likely to arise in various institutional contexts. If there are factors that make some institutions (e.g., markets, firms or courts) more or less likely to adopt or generate rules of exclusion, then neutrality on the question of institutional choice cannot be maintained.

Part 1: What is the Numerus Clausus?
The Law's Preference for Rules of Exclusion

"It is important to note… that the function of formal rules is to promote certain kinds of exchange *but not all exchange*."[18]

[18] North, I, IC, EP at 47(Emphasis added.).

Introduction

One of the primary difficulties in broaching the subject of the *numerus clausus* from within American law is, of course, that the doctrine does not officially exist in the common law system. However, the continuing reoccurrence of the principle against restraints on alienation, and the legal gymnastics that are undertaken to limit the availability of servitude-like restrictions, provides ample evidence that the common law has a heritage of robust suspicion towards the novel forms of property that private parties seek to create. Admittedly, the *numerus clausus* does not enjoy the secure and well-defined space that it occupies in civil systems. Nonetheless, it is a legal truism that even in the common law, property is significantly less customizable than contract. To locate the boundaries of contract, we generally look to the details of the agreement between the parties. To locate the boundaries of property, on the other hand, we generally begin by identifying the pre-existing legal 'type' into which the property fits. Only after we have done this do we look to the details of the boundaries that have been specified by the parties, and even then the question is often whether those privately set boundaries fit within the restrictions imposed by the pre-set legal categories. For example, if the 'property' in question carries with it less than the full range of uses ('full user'), we look to the law of servitudes to see whether the legal requirements of 'touch and concern', and other such limiting doctrines, are met. If not, then we are likely to cast the putative 'property' form into the rough seas of contract law, where the parties' intention to create a 'right against the world' is stymied by the requirement of privity.

Given the underlying ubiquity of the principle, the perspective adopted here is that, in the common law, the *numerus clausus* is in the nature of a 'judicial rule of thumb'.[19] That is to say, the principle operates as a form of judicial inertia, which limits the ability of private parties to inject new forms of property into the legal lexicon. It is critical to note that no assertion is made that there is a certain 'magic' number of forms that should be allowed, and indeed it is recognized that judicial and legislative recognition of new forms has, and continues to, occur.[20] Nonetheless, the principle is taken to be a significant limitation on the proliferation of property forms, even in common law systems. This accords with the view taken in the preexisting literature on the subject, which has adopted the civil law term '*numerus clausus*' to refer to the relative lack of 'private customizability' in the law of property, despite its lack of official status in the common law.

Current Theories

Because the *numerus clausus* is not an official doctrine in the common law, it is not surprising that disagreement has emerged about what the defining characteristics of the doctrine actually are. The two main theories – the 'third party information costs' theory constructed by Merrill & Smith, and the 'costs of verifying divided rights' rationale offered by Hansmann & Kraakman – each adopt their own version of the principle. In the Merrill & Smith framework, the *numerus clausus* is seen as creating a limited menu of property forms. When private parties seek to create a new property type, the law responds by trying to fit the new form into one of the old, familiar types. They divide this menu into 5 broad categories, and briefly list the property sub-types that have generally been allowed in each. The categories (and sub-types) they identify are: 1) estates in land (fee simple absolute, defeasible fee simple, fee tail, life estate, lease); 2) concurrent interests (tenancy in common, joint tenancy, marital property, trusts, condominiums/cooperatives/time-shares); 3)

[19] This convenient (and accurate) term is adopted from Merrill & Smith.
[20] The right of publicity is one of the often cited areas of such innovation. Similarly, Tulk v. Moxhay marked a watershed in the judicial recognition of new forms.

nonpossessory interests (easements, real covenants, equitable servitudes, profits); 4) interests in personal property (much more limited than estates in land - life estates are common, while other future interests, and equitable servitudes, are more limited); and 5) intellectual property (patents, copyrights, trademarks, trade secrets, right of publicity, misappropriation of information, and other *sui generis* forms).

Why is this limited menu of forms necessary? Merrill & Smith focus on the information costs that new property forms create for third parties (those who do not participate in transactions involving the property). They liken the number of property forms to words within a language: "The inventory of property rights can be analogized to the lexicon of a language, and the rules for combining property rights are like a language's grammar."[21] Limiting the number of property rights reduces the costs to those *outside* a transaction, because the effort they have to put into determining the parameters of any particular transaction is less than it would be if property were infinitely customizable. Because third parties only have to be aware of a fixed number of forms, the effort required to learn the lexicon, and to identify which particular property form is before them, is less than it would be if property forms were infinitely customizable. In effect, without the standardization of the *numerus clausus*, third parties would potentially have to 'learn something new' each time they came across a piece property. And even if the property in any one case was not a new form, they would have to investigate just to be sure. Standardized property bundles thus reduce the costs that third parties must incur to determine the specific attributes of any and all property that they encounter. Looking out at the universe of property rights, a third party would not know *ex ante* the attributes of any particular piece of property he comes across, and he would therefore have to incur costs in order to investigate and determine the boundaries, restrictions, etc., of each and every property right he comes into contact with.

A good way to think about the theory is as follows: consider what the proliferation of household cleaning products has done to the everyday shopping experience of the consumer. Several years ago, products used to be largely differentiated by a small number of large brand name offerings. And there were not only fewer brands, there were also fewer product types within each brand. This made shopping easy: you could quickly and easily identify your preferred brand, often by the general color of the packaging alone (orange in the case of Tide, blue in the case of Cheer, yellow in the case of Sunlight). Now when you look at a shelf of products, you are faced with a dizzying array of brands, and product features within each brand and product type. Even if you know *exactly* which product you want (you are not comparison shopping), it takes you longer to find it on the shelf. And you might have to stop and read the label twice just to make sure it is indeed the '2X Ultra Tide with Dawn StainScrubbers in Original Scent' and not the 'Tide with Febreeze Freshness Powder Spring & Renewal Scent'. The proliferation of products literally creates 'noise' which all consumers must now search through. In a similar manner, the ability to create novel forms of property (land, land subject to an easement, land subject to an unusual covenant, and so on) increases the information costs of third parties, even when they have no interest in transacting over that new form of property - and these costs are often not outweighed by any benefit internal to the transaction that created the new form.[22] The *numerus clausus* therefore seeks to balance the costs of identifying

[21] Merrill & Smith at 35.

[22] Merrill & Smith's theory actually contains two related assertions: first, that the common law's reliance on a limited number of standardized property forms reduces the information costs of third parties; and second, that there is a socially optimal level of standardization in the number of forms. The existence of this socially optimal degree of standardization arises because the ability of human beings to process the information they encounter, and to effectively communicate that information, is inherently limited. In the same way that language operates to facilitate communication by providing a menu of limited expressive options, the law facilitates commerce by providing a limited menu of property forms. Merrill & Smith at 38. For this reason, the theory is often called the 'communication and third party information costs'. The label 'third party information costs' is used in this work because the 'communication' in Merrill & Smith's theory does not

new property forms (measurement costs) and of providing notice of them (administration costs) with the 'frustration costs' that arise whenever legal rules prohibit private parties from creating new forms that meet their non-standard needs.

The second major economic theory of the *numerus clausus* principle is provided by Hansmann & Kraakman. Their 'costs of verifying divided rights' focuses, not on the ability to create new property forms, but rather on the ability to divide rights across tangible assets. They carefully review the conclusions of Merrill & Smith and offer an alternative explanation for the *numerus clausus*: that it is a concern with the costs of verifying the ownership and content of rights (and that it should therefore turn on the costs of supplying effective notice of this information). They take their cue from the civil law formulation of the *numerus clausus*, which operates to restrict the division of property rights *in assets* among more than one owner: "as a general rule, all property rights in an asset must be concentrated in the hands of a single owner rather than divided into partial rights shared among two or more persons."[23] In applying this civil law principle to the common law, Hansmann & Kraakman identify four main areas of divided ownership, and the types of property that each allows: "real property (cotenancy, servitudes, condominiums, mortgages, future interests); intellectual property (copyright, patents); security interests in personal property (chattel mortgages…); and the law of entities (partnerships, corporations, trusts). Outside these four fields, and the specific types of property rights for which each provides, the law makes it difficult to create partial property rights. Rather, all property rights in an asset are presumed to be held by one person."

In Hansmann & Kraakman's view, then, the *numerus clausus* is not a restriction on the menu of available property forms, but rather a regulation of the type and degree of notice that is required to divide rights in assets across more than one person. The purpose of this regulation is not to reduce the 'noise' that new property forms create for third parties, but rather to make it easier for those wanting to transact over assets to know who they should transact with (the verification problem). Although this verification problem is primarily one of knowing *who* to transact with, whenever rights in an asset are divided among more than one person, the issue of knowing exactly *what* rights each person owns inevitably arises as well. In their view, property law addresses this verification problem with an assumption of unitary ownership: after all, if all rights in a tangible asset are generally held by one person, then it is much easier for me to verify who I need to transact with, no matter what type of transaction I which to engage in. If I want to buy a painting that you 'own', and the law requires that ownership include the right to resell and asset, then I immediately know who to transact with. If ownership does *not* generally include the right to resell, on the other hand, I have to investigate further to determine if I need to transact with additional persons in order to purchase the painting. The law does, however, allow partitioning of rights in an asset across multiple persons: as long as adequate notice is given to all those who might be affected. The law therefore allows partitioning in those cases where the value of having a divided right is relatively large, and the cost of verifying that division is relatively low. This restrictive attitude towards dividing the ownership of the rights in an asset creates the *numerus clausus*.

---

always refer to communication among parties seeking to transact. The term can therefore be somewhat misleading, so it is not used here.

[23] Hansmann & Kraakman at 375. Of course, even the civil law allows some division of the rights in the asset, and these are noted: "exceptions include… cotenancy, servitudes on real property, mortgages on real property, and security interests in personal property. Partial property rights that do not conform to one of these specific exceptions are unenforceable." *Id.*

Why We Need a New Theory

Both of these theories represent significant contributions to the legal corpus. They are not only the foundational contributions to the literature on the *numerus clausus*, but they have done much to advance our understanding of both the practical contours and the underlying logic of the principle. Nonetheless, each fails to account for some persistent features of the law's resistance to the introduction of new property forms. In the case of the 'third party information costs' theory, the objections center primarily[24] around its focus on the costs to third parties (the 'noise' problem), rather than on person within the chain of asset transfer. The 'noise' problem only occurs in a select number of cases, cases which involve property that is actually highly customizable (because it involves product design or license terms). The 'noise' that raises the cost that third parties must incur to search through many property forms only occurs if you actually see or encounter the other property types. In most cases, this simply won't be the case: the third party will never even be aware of the other property types, he simply deals with the property over which he is seeking to transact.[25] And even if there are only a few property types, an individual will still have to check to see what type of property he is dealing with, so that is a constant cost. It is not clear, for example, why having an unusual easement in a deed is any more of an informational burden than having a type of easement selected from a limited list. There *are* cases in which the noise issue is actually a problem (the product design example is the best one), but these are cases in which the *numerus clausus* is actually at its weakest. Furthermore, the focus on third party costs is an uncomfortable fit with the structure and goals of private law, which generally seeks to adjudicate the rights as between individuals before the court. Although there is certainly a policy background that informs and shapes all of law, whether public or private, it would seem to be an odd thing for the defining feature of property law to have *nothing* to do with the actual owners and possessors of that property.

The 'costs of verifying divided rights' theory also greatly advances our understanding of the *numerus clausus*, and of related facets of legal theory. Objections to this theory revolve around the role that notice plays in regulating the entry of forms into the *numerus clausus*. Recall that this theory posits that the law allows rights in an asset to be divided when the cost of providing notice of that division is low, and the value of creating those divided rights is high. However, the law regularly *restricts* the division of rights in an asset even when the costs of verifying that division are extremely low, and the value of that division is extremely high. One recent example of this occurred in the *Quanta v. LG* case, in which the court refused to divide the rights in lawfully manufactured patented devices between the patent owner and the manufacturer of the device. As a result, a licensed manufacturer of chipsets and microprocessors could sell the devices to computer manufacturers, and the computer manufacturers did not need to pay both the component manufacturer (for the device), and the patent owner (for a license to use the device in assembling its computers). There was no question in this case that all parties knew of the patent owner's claim that it retained the right to license the devices (so there was no notice problem), and the value of the case was substantial (upwards of several million dollars for LG alone, not to mention the other patentees who would benefit from the ability to claim royalties from downstream transferees of patented devices). So, the law prevents the creation of divided rights even when everyone has notice of the division, and when a great deal of money is at stake.

An Alternate Definition: A Legal Preference for Rules of Exclusion

[24] Although notice also plays more of a role in Merrill & Smith's theory than in this one.
[25] An objection which is shared by H &K.

There is much in common between the 'third party information costs' and the 'costs of verifying divided rights' theories, and indeed, Hansmann & Kraakman see their explanation as complementary to that provided by Merrill & Smith. There are, however, several differences between them, and Hansmann & Kraakman carefully catalogue and analyze them.[26] For our purposes, the most relevant difference between them is that in the Merrill & Smith view, there is an asserted upper bound on the socially optimal number of property forms, so notice does not completely solve the problem that the *numerus clausus* is designed to address.[27] In Hansmann & Kraakman's account, this upper bound is dependent only on the actual cost of providing notice, and so technological and organizational capabilities provide the only limit to the number of forms that should emerge. To the extent that notice does more work in the 'costs of verifying divided rights' theory, the treatment of notice in this theory departs more from Hansmann & Kraakman's theory than from Merrill & Smith's theory. For the purposes of this work however, the similarities between the two theories are more important than the differences. The two similarities of particular importance are: 1) they both focus on the impact that novel property forms have on third parties, rather than on parties within the chain of privity[28] (the third party issue); and 2) notice plays a pivotal role both theories in determining the need for the *numerus clausus* (the notice issue).[29] The theory offered in this work takes these two aspects of similarity as a starting point, constructing a theory which: 1) focuses on the impact of novel property forms on individuals *within* the chain of title; and 2) which does *not* depend on the costs of providing notice of these forms. The choice of these two particular features is deliberate, because the *numerus clausus* is actually concerned with what goes on between the parties to the transaction (this will be shown below and in Part 2), and because it resists making the *numerus clausus* turn on notice alone (as we have already seen from the LG case).

To begin the construction of this new theory, we need a new definition of what the *numerus clausus* is. It is important to note at this point that although each theory of the *numerus clausus* has its own unique definition of the principle, these definitions are *not* inconsistent with each other. Indeed, each definition can be restated in terms of the other. So, for example, the law's limit on divided rights (Hansmann & Kraakman) also has the effect of producing fewer novel property forms (Merrill & Smith). However, each definition focuses on a different set of costs, and so each leads to a different degree of robustness for the principle. In Hansmann & Kraakman's theory, the *numerus*

---

[26] [Insert list here.]

[27] Merrill & Smith cite this upper bound as being produced by the intractable features of human 'communication'. It is not completely clear why notice does not dispose of this upper bound, however. Merrill & Smith deal with this objection by stating that notice is always costly to process: "Making the running of a fancy depend solely on the original parties' intent and on notice - even recorded notice - to subsequent parties acquiring property assumes that notice is the most cost-effective method to minimize third-party information costs. But notice of idiosyncratic property rights is costly to process." However, it is not clear why notice should not serve to reduce such costs, particularly when we consider those situations in which the terms are embodied in the 'product' itself, and so require no processing at all. See Part [].

[28] "We agree, in particular, that third-party information costs are central to the law's regulation of property rights." Hansmann & Kraakman at 374. It is not completely clear that Hansmann & Kraakman always use the same definition of 'third parties' as that used by Merrill & Smith. Merrill & Smith use the term to refer to those outside the chain of transfer of any particular piece of property, while Hansmann & Kraakman most frequently use the term to apply to third party transferees of an asset, though they also refer to true third parties who need to interpret the ownership and content of rights, such as the judiciary. However, because the Merrill & Smith definition is more restrictive (and thus more difficult to respond to), Hansmann & Kraakman's assertion of equivalence will be taken at face value. To the extent that there is a difference (and it seems that there must be, because if you are seeking to transact over an asset, you are by definition a first party, even if just a prospective one), the only consequence for this theory is that its focus on first parties represents a departure from the Merrill & Smith theory, while its treatment of notice departs from both theories.

[29] If third parties can be cheaply, quickly and accurately notified of unusual forms of property, notice can substitute for standardization. Despite Smith and Merrill's assertion that notice 'does not cure all' within their framework, they do explicitly recognize that decreasing costs of notice will lessen the need for the doctrine.

clausus is needed because verifying the ownership of divided rights is costly. When cheap and effective notice reduces the cost of this verification, therefore, the *numerus clausus* is no longer necessary. In Merrill & Smith's framework, on the other hand, more property forms will always increase the information processing costs of third parties, and notice does not completely reduce these costs. In this work, the *numerus clausus* is seen as the law's way of implementing a preference for rules of exclusion (which, by definition, limits the proliferation of governance rules). What exactly is a 'rule of exclusion'? When an individual operates under a rule of exclusion with respect to an asset, he can 1) unilaterally determine the use of that asset 2) without paying another party for the right to do so. The most convenient short-hand term for exclusion is perhaps 'full user',[30] a term which accurately describes the concept (you get a full bundle of legal privileges with your property), but which has never been extensively used in the literature. A rule of exclusion is best understood in terms of what we normally consider to be the 'right to exclude' in property law: your right to exclude generally carries with a full set of legal privileges (use), powers (transfer) and immunities (income). As a property owner, you alone determine what uses are to be made of your property. Under a rule of exclusion, in other words, full use generally travels with permanent possession of an asset.[31] In governance rules, on the other hand, uses are either specified and/or they must be paid for (there is either no 'right' to unilaterally determine a use, or, if there is, that use must be paid for).[32] Under a governance rule, therefore, rights of use are at least partially separated from the right to possess an asset.[33] Servitudes, which separate possession of an asset from certain use rights in an asset, are the canonical example of governance rules. The law's antipathy to servitudes is thus an aversion to governance rules.

The governance/exclusion terminology, which was pioneered by Henry Smith, may be unfamiliar to most readers, and thus may be somewhat opaque. However, it is nonetheless superior to all of its near substitutes, of which there are several. [Why 'license or sale', 'contract or property', is not sufficient will be added here.] Now that we have established that the *numerus clausus* operates to limit the proliferation of governance rules in property law, we need to ask *why* it does this. Can we make sense of this deep-seated judicial impulse?

The Legal Preference for Exclusion vs. Governance: Why Prior in Time is not Always Prior in Right

To begin our examination of why the law displays a preference for rules of exclusion as opposed to governance rules, it may be helpful to restate the issue in more familiar terms. Why does the law overwhelmingly disfavor use restrictions when a tangible permanently changes hands, and why, after a 'sale', is a previous owner no longer entitled to increases in an item's value? To answer that these characteristics embody the very notion of 'sale' or 'property' provides no answer at all.

---

[30] One which brings to the fore the similarity to the Hansmann & Kraakman definition.

[31] 'Permanent' is used as a qualifier here to distinguish changes in possession that are merely temporary (e.g., rental of a chattel).

[32] Note that the requirement to pay for a use is equivalent to not having the right to a property's income stream. It should be noted that the definitions of exclusion and governance used here are similar, but not identical, to those used by Henry Smith. In his definition, governance rules are characterized by specified uses. Although uses are often specified in governance rules, they need not be. An example would be the case in which the right to control and the right to income were completely separated from each other. That is to say, there is a completely open-ended, but fully priced, liability rule.

[33] 'Permanent' is not used as a qualifier here because governance rules can be created when an asset has changed hands either permanently or temporarily. The *numerus clausus* seeks to regulate only those governance rules that accompany the permanent transfer of an asset.

Indeed, we could imagine a legal world in which resale/reuse restrictions and royalties were the rule, rather than the exception: property scholars will recognize this as a world in which bailment-like arrangements dominated.[34] Circular reasoning (new property forms cannot be legally recognized because they do not fit within pre-existing legal categories) has been one of the major weaknesses of the justifications found in the case law,[35] and has been used to great effect by those arguing for a relaxation, or even eradication, of restrictions on the creation of servitudes and novel licensing schemes, for example.[36] The search for a rationale in the shifting sands of the common law is made even more precarious by the fact that the 'single reward', 'full value', or 'just compensation' concept ('sale' of 'property' requires that changes in value accrue to the transferee rather than the transferor) and the 'restraints on alienation' principle ('sale' of 'property' requires that there be no restrictions on use) are being questioned ever more frequently by the courts themselves, as they realize the question-begging quality of these legal truisms. The question comes down to this: why do rights have to transfer at all, if the initial owner doesn't wish them to do so? Consider a garden variety item of personal property. If I, as an initial owner (or producer), own all rights in a camera, for example, why can't I retain some of my rights, while transferring others? I might want to retain the right to resell or lend the camera, as this will enable me to sell a new camera to each person who wants one.[37] When I owned (made) the entire camera, the right to resell or lend it was certainly in my bundle of rights - so why can't I retain it after physical transfer of the 'thing'? Why, in other words, should the law force an initial owner to transfer what used to be hers? The most comprehensive articulation of this quandary, and its implications for contract and property law, can be found in a series of works of Richard Epstein.[38] But the issue has also been taken up in the recent work of other scholars: "start with the premise that the power to restrict use is intrinsic to the power to transfer (or not), which is fundamental to the concept of property rights. Exercising that power means, of course, that prior owners can limit the rights of subsequent owners; this simply reflects the temporal order of property rights: *first owners determine the rights of second owners*. We could reverse the sequence, but it isn't easy to explain why the second owner's right to unburdened use of property

[34] Replacing sale with a series of bailment-like arrangements might sound improbable, until we consider that the software industry attempted this very thing. See Part 5. Two existing examples of use restrictions and resale royalties are analyzed by Hansmann & Kraakman in their paper: the right of integrity, and the resale royalty applied to works of art in some continental jurisdictions. The right of integrity gives an artist the right to prevent the destruction or modification of his work, even after a copy of it has been sold, and is owned by another. The resale royalty requires that an artist receive a share of the proceeds every time his work is sold to another buyer. Although these examples are both from the intellectual property field, Hansmann & Kraakman correctly note that they do have counterparts in the law of tangible property. In the case of the right of integrity, for example, government mandated restrictions on the modification of historical properties, and environmental preservation servitudes, are counterparts in the realm of real property that are becoming much more common. Outside of the intellectual property context, resale royalties are much less common, as they lack the paternalistic justification of allowing maturing artists to share in the increases in value of their works as their reputations develop.

[35] A close reading of Kepell v. Bailey, for example, reveals that the case had two alternate grounds of decision. The first was that the covenant in issue directly contravened the local Canal Act; while the other was that the covenant was "unknown to the principles of law", and did not comply with the various limiting doctrines, such as privity of estate, that were required for covenants to run with notice. Similarly, the foundational decision in the intellectual property context, Bobbs-Merrill, merely stated, without explanation, that the imposition of a servitude-like restriction on a book was outside the Congressional copyright grant. Hence, reading the tea leaves of the common law in this matter is a largely a fruitless enterprise, whether the leaves are found in the servitude cases, the standard form contract cases, or the 'restraints on alienation' cases.

[36] E.g., Robinson, Susan French, Epstein.

[37] "I just returned a digital camera because it had a hardware shrinkwrap license… They said that anyone in my household could use it but I couldn't lend it to anyone else." The license in this example did explicitly allow purchasing the item as a gift, and resale was not specifically mentioned (though why it would seek to restrict lending but not resale is not immediately apparent).

[38] Possession as the Root of Title, Notice and Freedom, etc.

should trump the first owner's rights to burden the property."[39] This is the charge that any theory of the *numerus clausus* must answer.

The Function of Rules of Exclusion: Allowing Asset Transferees to Breach Governance Rules

At first glance, the question of why the law might force a reallocation of rights from seller A to buyer B seems almost unanswerable. But if we first envision the problem as a three-person instead of a two-person problem, we might see things in a different light. Consider the following scenario: Property owner A owns a plot of forested land, on which he currently runs a paint ball business. A has signed an agreement with C, that C and his heirs can play paint ball on the land every Wednesday until 2025. B buys the land from A. In the deed of sale, no explicit mention is made of the paint ball business, or of the use to be made of the land, but B is told about A's agreement with C. B opens his own paint ball business, and C shows up on Wednesday, demanding to play. Should B have to honor the agreement? There would seem to be 3 potential legal responses to this state of affairs. 1) B is allowed to ignore the agreement between A and C because although he had notice, he did not specifically agree to abide by the agreement. C can still sue A for breach. Under this approach, both B and all downstream transferees operate under a rule of exclusion. 2) B must abide by the agreement with C because he had notice, but can circumvent the agreement by selling the property to D without giving D notice.[40] C can still sue A, A can (possibly) sue B, but no one can sue D. Under this approach, B operates under a governance rule, but any downstream transferees operate under a rule of exclusion. 3) B must abide by the agreement with C if he has notice, and so must any downstream transferees of the land if they have notice. When notice is expensive or difficult to provide, this scenario looks much like that in response #2. If notice can be permanently affixed to or associated with the land in some way, however, then all transferees of the land have notice and operate under a governance rule. These results are summarized below.

Table 4: Legal Approaches to the Three-Person Problem

Scenario	Legal Regime	B Operates Under What Type of Rule?	Downstream Transferees Operate Under What Type of Rule?
1	Contract	Exclusion	Exclusion
2	Bilateral Notice	Governance	Exclusion
3	Multilateral Notice	Governance	Governance (with notice only)

[39] Robinson at 1462. (Emphasis added.)
[40] A strategy which might not seem to benefit B. But if D were a wife or a relative of B, this might be a sensible strategy.

How should the law choose among these responses? Since B had notice of the agreement in each scenario, it seems likely that B paid a lower price for the land than he would have if no such agreement existed, so equity concerns would point towards giving C an enforceable right. Indeed, if the ultimate use of the land is as a paint ball venue, it seems wasteful to require C to engage in some combination of suing A and renegotiating with B, just to allow him to do what he had already paid to do.

But suppose the situation were a bit different. B is about to start his own paint ball business when he reads about a huge subdivision being built in the area. B has an epiphany: 'What this town needs is a Mall.' He immediately starts cutting down trees. C shows up, demanding that he be allowed to use the land as per his agreement with A. (Assume that B can either build a mall, or allow people to play paint ball, but not both.) What should the law do in this scenario? Suppose that B offers to pay C to relinquish his right to use the land. If C agrees, then B will likely end up making a payment to C roughly equal to the discount he got in buying the land from A. In this case, all parties end up in roughly the same position they would have been in if there never had been any agreement between A and C, and if A had simply sold the land outright to B. Thus, there seems to be no harm in legally recognizing C's right, whether that right is seen as merely contractual, or as a property right that runs with the land. But suppose that when B offers to pay C to relinquish his right, C plays hard ball. If C simply demands more money than he paid for the right in the first place, there is little harm done. B ends up paying a higher price in total for the right to use the land however he likes, C ends up with a little more money, and A ends up in the same position. There is no absolutely economic reason to be concerned with the different distribution of funds between B and C in these two cases. But suppose that C does not just play hard ball. Suppose that C refuses to be bought out at any price. Perhaps C simply dislikes B, or perhaps C has a hidden agenda because he loves trees and hates malls. In other words, perhaps C's preferences simply do not respond, for whatever reason, to price. In this scenario, giving C a right can affect the ultimate use of the land. Furthermore, there is a potentially huge difference between giving C a right that is 'contract-like' (one that can be breached and paid for in damages[41]); and a right that is 'property-like' (one that can force actual compliance with an injunction).

Assume for the sake of argument that the use of the land as a mall rather than as a paint ball venue really is a 'higher-value' use. If buy-out is relatively simple, then both 1) allowing B to ignore the agreement between A and C (whereupon C sues A), or 2) requiring B to pay C to relinquish his right, results in a form of 'efficient breach'. If C *can* be bought out, in other words, the second scenario is roughly equivalent to allowing B to breach and pay damages, even if C's right is equitable (i.e., injunctive) in nature. In the first case, B's actions cause A to breach; in the second case, B breaches but pays C for the right to do so. Although there is perhaps a moral difference between these two scenarios, and there is certainly a distributional difference, there is no economic difference. If C *cannot* be bought out, on the other hand, there is a pivotal difference between the first scenario (in which B operates under a rule of exclusion) and the second (in which B operates under a governance rule). In the first situation, C gets damages from A, and B builds a mall. In the second situation, B must allow C to play paint ball, and no mall is built. The ultimate use of the land is impacted, and there is an economic difference in the two scenarios because production, as opposed to distribution, is affected.

[41] But note that the measure of damages (expectation versus 'disgorgement', for example) becomes critical here, as does the availability of specific performance. Here, disgorgement is used in the looser sense of 'my profit versus your loss', rather than in the stricter sense of a purely restitutionary measure.

Now suppose that we have a two-person problem, instead of a three-person problem. Suppose that A says to B: "I want to sell you all rights in this land, save for one - I want to retain the right to use the land to play paint ball every Wednesday," and that B agrees. How should the law interpret such an arrangement? Once again, there seem to be (roughly) three legal responses. 1) The law might view this as an unsuccessful attempt to create a new property form. If the legal system in question contains a *numerus clausus* restriction, then the law can simply refuse to recognize the arrangement because it does not fit within a pre-existing category. B, and any downstream transferees, then operate under a rule of exclusion. 1a) The law might interpret this situation as 'merely contractual'. In that case, B can breach the paint ball covenant, and pay A damages. Any downstream transferees of the land can buy the land unencumbered by B's original promise to A – *even if that promise is what motivated A to sell the land in the first place*. 2) To mitigate the harsh effect of the *numerus clausus* (and of contract law!), the law might take the middle path of equity, and allow the paint ball covenant to be enforced by injunction as long as notice is given. In this case, both B, and any downstream transferees with notice, must abide by the covenant. 3) Finally, the law could consider this to be a permissible property transaction. This could occur in legal systems without a *numerus clausus*, or, in legal systems with a *numerus clausus*, if the form in question was a permissible entrant into pre-existing categories. In this scenario, two new forms of property are created. In addition to 'land in fee simple', we now have 'land subject to a paint ball covenant'[42] and 'benefit of a paint ball covenant'. A has the right to use the land on Wednesday's to play paint ball, and B has all other uses (to the extent that they don't interfere with A's right). In order to use the land inconsistently with A's retained rights, B must now seek A's permission (and pay him for that permission). B operates under a governance rule as regards his land, rather than a rule of exclusion. In addition, if B transfers his 'land subject to a paint ball covenant' to C, C must also get A's permission before using the land inconsistently with A's right. And so on down the chain of transfer. These possibilities are summarized below.

Table 5: Legal Approaches to the Two-Person Problem

Scenario	Legal Regime	B Operates Under What Type of Rule?	Downstream Transferees Operate Under What Type of Rule?
1	*Numerus Clausus*	Exclusion	Exclusion
1a	Contract	Governance (breach with expectation damages only)	Exclusion
2	Property Customization Allowed with Notice	Governance	Governance (with notice only)
3	Full Property Customization Allowed	Governance	Governance

[42] It is not clear whether this would be considered an easement or a covenant. To the extent that A can simply do something on someone else's land, it looks like an easement. But to the extent that that 'something' is of a type that actually prevents B from doing something on his own land, it looks more like a covenant. For the purposes of this hypothetical, it is not important whether this would be considered to be a covenant or an easement, since we are ignoring all the complexities (privity, touch and concern, etc.) that go into the law's determination of what constitutes an allowable covenant, easement or servitude.

Comparing Scenario 1 in Tables 4 and 5, we can see that there is an unexpected correspondence between the function of breach in contract law, and the operation of the *numerus clausus*. Both operate to replace a rule of governance with a rule of exclusion vis-à-vis downstream transferees. Of course, in the contract case, the initial promisor must pay expectation damages when this occurs (this can be seen in Table 5, Scenario 1a), and in this sense, contract is rightly considered to be a governance rule. But contract, at least in its Holmesian conception, is a very special type of governance rule. It is special in two respects: first, it allows for breach (instead of performance) and the payment of expectation damages (instead of a measure that requires the disgorgement of the profits of the breaching party); second, it attaches only upon agreement, rather than by mere notice. Both of these characteristics work to replace a governance rule with a rule of exclusion: the first, by allowing efficient breach by the original promisor, and by maintaining the incentive for that breach with expectation damages; the second, by making it difficult to bind downstream parties to the original promise. Of course, the line between actual agreement and mere notice is often hard to draw, particularly when we consider the mass market context. Without the standard form contract, which is neither negotiated, nor (most often) read, modern commerce would be impossible. The relaxation of contract formalities in the mass market context makes it tempting to view 'offer and acceptance' as nothing more than a legal fiction. But 'contract by notice' is not really contract at all, as it is potentially binding on the whole world. The cheaper and easier it is to provide notice, and the more notice is taken to be constructive rather than actual, the more contract begins to look like property.[43] But not just any type of property – for what we have then is property that has freed itself from the restrictions of both contract *and* property law. When we combine the customizability of contract with the binding power of property, in other words, we get a potentially endless variety of infinitely customizable servitudes. In such a world, rules of exclusion become replaced by governance rules (we move from Scenario 1/1a to Scenario 3 in Table 5). Concerns about equity (B should be bound by that restriction because he had notice of it) also move us from exclusion into governance. It is easy to see why notice has been such a huge battleground in private law: if we take either 'the equity of notice' or the 'agreement is a fiction' objections too far, then the legal conceptions of *both* contract and property are completely overwhelmed. Like Castor and Pollux, they rise and fall together.

We can see from the above discussion that contract (in its Holmesian formulation) and property (in its legal, rather than equitable, incarnation) are not really opposing legal forms. Instead, they are simply two ways to mediate the re-occurring conflict between rules of exclusion and governance rules. In other words, contract and property allow two different forms of breach. Although I will eventually assert that *both* of these forms of breach are efficient, for now let's simply say that there is some sort of breach in both cases (which may or may not be efficient). As illustrated by the three-person and two-person problems examined above, contract and property both operate to replace governance rules with rules of exclusion. Whenever a rule of exclusion replaces a governance rule, breach occurs because downstream transferees can now seek out new uses of assets without the permission of upstream parties (whether these be parties to a contract or former owners of assets). Viewing both contract and property as enabling a form of breach should not be surprising, because governance rules can arise either from the live hand of contract, or the dead hand of property. In the former case, a bilateral agreement is passed

[43] Strictly speaking, either specific performance or full disgorgement of profits is also required to complete the equivalence.

down via a series of sequential promises to bind all assignees; in the latter, a new property form requires compliance without any agreement at all. The damage levels for the two types of breach are, of course, different. In the contract case, the breach arises from a specific bilateral promise, and the damage level therefore reflects the expected outcome of that promise (expectation damages). In the property case, the breach arises from a pre-existing, multilateral legal obligation which was not specifically agreed to, and so the damage level is set at zero[44] (even though the actual 'damages' of the former property owner are greater than zero). What is this 'pre-existing multilateral legal obligation'? It is nothing other than property ownership. Consider once again the two-person scenario discussed above. Recall that A used to own the entire lot of land and all rights in it, and at no time did he consent to give up his right to use the land on Wednesdays to play paint ball. When A owned the entire lot of land, there is no question that the rest of the world had a legal obligation not to interfere with his use of that land (no one could prevent him from using that land to play paint ball on Wednesdays). Furthermore, there is no question that the rest of the world was bound by A's property right even though no one specifically agreed to be so bound. (That is, after all, what it means to have property – you don't have to get someone's permission to use it.[45]) It is clear that A did not transfer all of his rights in the land in his arrangement with B; nonetheless, the *numerus clausus* steps in and allows B, and any downstream transferees, to 'breach' A's retained right to use his land on Wednesdays. Furthermore, they can do this without having to pay damages. Whenever the law refuses to enforce a servitude, or whenever the law transforms an attempted license into an outright sale, efficient breach occurs. Just as contract allows the breach of a promise, sale allows the breach of a license,[46] and property (full user) allows the breach of a servitude. In other words, *rules of exclusion allow the breach of governance rules*.

If this proposition seems difficult to accept, perform this simple experiment. First, search your office or your home for an item which has a use restriction printed directly on it. These are more common than you think, though you might have to look closely for the tiny print. It might be a DVD which is licensed only for viewing only in the United States (I rented one of these recently from Blockbuster), or a directory from a professional association which prohibits its resale or transfer. It might even be a plant which prohibits asexual reproduction (though in this case it would still have to have the tag attached, as they haven't figured out how to print the text directly on the plant yet[47]). If you are a tinkerer and a pack rat, you might go into your garage and look for some old vacuum tubes which are licensed for use in car chargers, but not in radios. Whatever item you find, hold it in your hands, read the text carefully, and then imagine doing exactly what the text prohibits you from doing. If you can, it would be better if you could actually perform the prohibited action. Now step and back ask yourself: What have I just done? The answer is quite clear. You have breached a governance rule. Why have you done this? Most likely, because you consider this item to be your property, and you think that you should be able to use or dispose of it however you want. Luckily, the physicality of the item enabled you to do

[44] Actually, the lack of specific agreement is only the first of two reasons why the damage level is set at zero. The second reason is discussed later, in Part 2.

[45] Or pay them for that permission, even if it is automatically given, as in a liability rule scenario.

[46] Even though licenses are generally parsed as contractual in US law, the term 'license' in its proper sense refers to a unilateral abrogation of rights which can be revoked at will by the licensor. In this sense, no actual agreement is required for a license to be granted. Of course, the more complex 'licenses' become, they more they begin to look like contracts which must be agreed to by both parties, rather than mere unilateral permissions which can be revoked at will.

[47] Though, of course, Monsanto has prevented how to produce single generation seeds.

this (keep this in mind because we will return to this point much later, in Part 3). Breach was *de facto*. But if the law allows you to ignore these notices (whether they are conceptualized as a failed contract or a failed property innovation) they actively allow you to breach a governance rule. Breach in such cases is *de jure*. So the question arises, why does the law countenance such blatant noncompliance? Such is the business of Part 2.

Part 2: User Innovation – When Breaching Governance Rules is Efficient

"[A] growing body of empirical work shows that users are the first to develop many and perhaps most new industrial and consumer products."[48]

[48] Von Hippel.

In this Part, we will consider the general question of when the breach of a servitude (or other governance rule) is efficient. The answer to this has already been briefly alluded to in our '2-person' and '3-person' problems in Part 1: breach will be efficient whenever it allows for a higher-value use of the burdened asset. That is to say, breaching a governance rule will be efficient when possessors of assets use those assets in the process of innovation. The focus on user innovation gives rise to the first major departure from previous theories, as it brings us squarely within the chain of privity of the asset. Third party concerns play no role in the user innovation story, and so they play no role in the efficient breach theory of the *numerus clausus*. The focus on user innovation also marks a point of departure from the related literature on servitudes. That literature has identified 'the problem of the future' as an animating concern in the law's restriction on the creation of novel servitudes. The idea is that, just as the rule of perpetuities is concerned that the dead hand of property owners will restrict the freedom of future generations, servitude law is concerned that present property owners will not be able to accurately predict the needs of future owners. This has some similarity to the perspective that will be developed here: it obviously presents us with a concern about what happens within the chain of transfer, and it expresses a general unease with the control of future uses of assets. However, it is a concern which is not realized until some far away, unknown time in the future: it is the sheer length of time that causes the waning of the needs of the property owner (who is now dead), the waxing of the needs of future owners, and the dissimilarity between them. Because of this, these theories support the expiration of servitudes (or some strengthening of the changed circumstances doctrine), rather than an outright prohibition on their creation. After all, if the future truly is the problem, then 'they' can best sort it out when they get there. Indeed, prohibiting the creation of servitudes under this rationale would present us with an *exacerbated* version of the servitude problem in reverse (i.e., the unknown needs of the future would be restricting the needs of the present, rather than the known needs of the present restriction the unknown needs of the future). Given the choice between these two scenarios, the known and actual good is to be preferred over that which is unknown and hypothetical. There is another difficulty as well. Justifying servitude law (and by extension, the *numerus clausus*) on the unknown needs of distant future owners presents us with the difficult proposition of weighing the hypothetical needs of the unknown and the unborn with the very real and pressing claims of present owners who know their needs and wants (interests that they might want to protect via the creation of novel servitudes or other governance rules). The device of private property is supposed to alleviate this problem, by providing the owner with an incentive to weigh both the present and future income streams of his property in an effort to maximize its present value.[49] But it is not at all clear that property owners either have the will[50] or the way[51] to perform this calculation, even if private owners are indeed better than communal ownership in this regard.[52] But focusing on user innovation gets us out of this difficulty: it is not some unknown owner we are concerned with, but the actual possessor of the asset; it is not some unknown future time, but the here and now that is implicated with every transfer (resale) of the asset.

Now here is where things get tricky. It is indisputably true that the law does not, in most cases, make an inquiry into whether the new use to be made of a burdened asset is of greater

[49] See, e.g., Demsetz, Towards a Theory
[50] Their time discounting may be too great.
[51] They may simply not be able to predict future contingencies
[52] Add memorable quote about conversations with the unborn.

value than that which existed under the servitude.[53] In most cases, the law does not even ask whether any new use has, in fact, been made or found (as when it get rids of resale or price restrictions, for example). However, this fact alone is not fatal to the idea of efficient breach in property law: after all, contract law allows breach as a general matter,[54] regardless of whether it turns out to have been efficient in any particular case or not. But the underlying method in the madness that is property law goes deeper than this, because a rule of exclusion, although it allows efficient breach, is not fully exhausted by that concept. Whenever a governance rule is replaced by a rule of exclusion, four possibilities present themselves: 1) there is a successful search for a new use (efficient breach, intentional innovation); 2) there is no directed search, but freedom of use allows a fortuitous accident which reveals a new use (efficient breach, accidental innovation); 3) there is a search for a new use that fails (failed innovation); and 4) there is no search at all, and the breaching party receives a windfall (no innovation). Of these various possibilities, some are productive and some are not, but it is impossible to predict in advance which situation will transpire.[55] Therefore, in order to create the conditions under which successful searches for new uses can occur, property law must also allow for those cases in which no new uses will be found, or even looked for. The way it does this, of course, is by vesting a new property right in the downstream party.

As we will see, in some ways the terms 'search' and 'new use' inadequately convey the full range of innovative activity with which we are concerned. Sometimes users will engage in a conscious, directed innovative process; at other times, innovation will be the mere by-product of repeated use, or of accident. 'New use' can be similarly deceptive. Sometimes, user innovations will literally involve a new use of an existing asset. Other times, however, the innovation will be a modification of an existing asset, or the creation of a completely new asset – spurred on by the problems revealed by the use of existing assets. In still other cases, user innovation will involve the creation of complementary assets – assets which are specifically designed to be used in conjunction with assets purchased in the market (and which may or may not require some degree of modification of the market asset). We turn now to take a closer look at some actual case studies, first of intentional and then of accidental innovation, which illustrate how and why these various types of user innovations occur.

Intentional Innovation

The case studies in this section will be divided into two rough categories: problem solving and problem finding. These broad categories correspond to the distinctions found in the psychological literature on creativity. Just as importantly, they provide useful guides as to why user innovation has always been, and will continue to be, critical in moving assets to their highest-value uses: because user innovation occurs as a by-product of repeated asset use, users will have an advantage in both *finding* problems and in *solving* them. Although we may not be accustomed to thinking about problem finding and problem solving separately, they are two quite distinct creative events. Indeed, some individuals (including such erudite ones as the theoretical physicist, Freeman Dyson) believe that they have a talent for one activity, but not the other. Problem finding and problem solving can occur almost simultaneously, or they can be separated

[53] The doctrine of changed circumstances might, however, be considered to be an example of this type of calculus.
[54] Except in those exceptional cases where specific performance is warranted.
[55] We will discuss why this is so in more detail later in this Part, and in Part 3.

in time – which may be measured in decades, or even centuries. In the case of the first computer assisted surgical robot (AESOP), for example the problem (the need for better control of the viewing camera during laparoscopic surgery) and the solution (using a robot to control the camera instead of a human being) were found simultaneously. In the case of determining longitude while at sea, the problem (how to accurately determine longitude at sea) was known centuries before a solution presented itself (the marine chronometer had its first sea trial in 1736). Sometimes, the same individual will both find and solve a problem; other times, these tasks will be completed by different individuals. Sometimes, and perhaps most counter-intuitively, a problem will be solved even before the problem is found. At such times, a 'solution' exists for a period of time as a mere curiosity, with no known practical application. Only later does the 'problem' present itself, often in a completely different context. Such was the case with Post-It notes, for example. The adhesive for this now ubiquitous paper product was the result of a failed attempt to create an improved version of scotch tape. Its inventor, Spence Silver, was employed by the 3M Corporation. He was fascinated by the unusual adhesive he had found – it would not bond securely because it broke up into small spheres the size of paper fibers. Despite his scientific savvy and his continuing interest in finding an application for the adhesive (he spoke about it every chance he got), Spence Silver was not the one who came up with the idea of 'Post-It' notes. This flash of insight, this act of 'problem finding', came to us courtesy Art Fry, an employee in 3M's product development division. Fry had heard about Silver's unusual adhesive in a corporate seminar, but did not think about its possible applications until he was in church choir, and his makeshift bookmarks (scraps of paper) once again fell out of their assigned places as he turned the pages of his hymnal. It was then that he thought about Silver's adhesive – and how it might provide a 'permanently temporary' way to bookmark his pages. In this case, the 'problem' (flagging pages in a document) was not found until roughly five years after the 'solution' (the unusual adhesive).

Before we move on to a closer look at some specific examples of user innovation, a few words about terminology are in order. First, it is important to note that the terms 'creativity' and 'innovation' are used interchangeably to refer to the bringing forth of something new – whether it be an idea, product, use or modification. Neither of these terms corresponds in any way to the terms 'originality' or 'invention' as used in copyright and patent law – i.e., no judgment is implied about the degree or type of creativity or innovation in any particular case. The focus on newness rather than eligibility for IP protection is consistent with the well-known fact that much IP actually has little, if any value. Most patents, for example, have no economic value at all. In contrast, much sub-patentable and non-copyrightable innovation is extremely valuable indeed: databases, fashion, myriad incremental product improvements – they all generate significant revenue while falling under the bar of IP protection. Furthermore, some studies have shown that 'low-tech' innovations – mundane improvements in existing products, even in business models[56] – is the real life blood of many industries, even countries: "[N]ew products, processes and ways of organizing frequently evolve gradually in ways that have little or nothing to do with what we usually associate with R&D. Innovations, in other words, are often non-research-based and non-science-based."[57] Some have gone so far as to suggest that the current focus on 'high-tech' innovation in policy circles is counter-productive and even harmful. In other words, when it

[56] For an example of how the law's refusal to allow use restrictions helped to transform a business model, and even an industry, see *Infra*, '*The'Mere' Right to Rent/Resell: Is It Non-Innovative?*'.
[57] Jacobson

comes to economic success, information production (the generation of incremental improvements, new uses, and new ideas) may often be more important than IP production.

Second, the term 'innovation' is used in this work to refer to the originating act of creation or invention, rather than the act of bringing a finished product to market. Although some scholars are of the opinion that 'innovation' should only be used to refer to commercialization, and 'invention' should only be used to refer to the act of creation, other scholars use the terms synonymously. In particular, there is a growing literature on 'user innovation' which studies, analyses and documents the innovative activities of users of assets. In this literature, a 'user' is defined as anyone (including a firm, an individual consumer or a professional) who extracts value from an asset by using it, rather than by selling it.[58] When a user innovates, therefore, a new product is often created, but it is not created to be sold on the market. Some users may later decide to market their innovations to manufacturers, and some users may even become manufacturers themselves,[59] but the initial impetus in user innovation is always *use* rather than sale. This is in marked contrast to innovation by manufacturers, in which a new product is created for the sole purpose of being sold on the market. Of course, user innovation is not just important because it provides individual users with isolated solutions to their individual problems, and moves specific assets to higher-value uses. More importantly, user innovation often bleeds over into user communities, which adopt the new practice or product (and often refine it further), and into the marketplace, as manufacturers adopt the innovation and themselves begin to manufacture it. This type of innovation, in which manufacturers mine the innovations of consumers, is far more prevalent than once thought. In fact, "a growing body of empirical work shows that *users are the first to develop many and perhaps most new industrial and consumer products*."[60] (Emphasis added.) Of course, this empirical work measures both the innovations developed by firms for use in house, and the innovations of individual consumers developed for personal (e.g., a sports enthusiast modifies his equipment to improve his performance) or professional use (e.g., an electrician or plumber modifies a tool to make his job easier). However, businesses are learning that they can adopt and build on the actual innovations of their customers (rather than just their 'product suggestions'[61]), and some niche industries are built almost entirely on such innovation (as is the case of high-performance windsurfing equipment, discussed in the next section).

Problem Solving

We will begin by examining user innovation in four different areas: high performance wind surfing, skateboarding, computer software and computer games. These have been chosen because they have been the subject of detailed academic study, and because users of these products are particularly active, not only in innovating, but also in sharing their innovations with others. This sharing of innovations has two salutary effects. First, it enables other users to build on and further refine the innovations through their own use of them. Second, it also increases the likelihood that manufacturers (if they take the time to look for user innovations) will become of aware of the innovations and begin to sell them in the marketplace. This increases the economic

[58] Hippel, DI at 3.
[59] This often happens in the case of sporting goods, for example, DI at 127.
[60] Hippel, at 2.
[61] Note that these are two very different things – von Hippel, Christensen.

impact of the innovation, and disseminates the innovation to less sophisticated users (i.e., users who prefer to buy a finished product in the marketplace, rather than modify something they already own). It is no accident that the examples chosen are from the fields of individual sports and computer programming. Both of these fields share several characteristics which make them particularly amenable to user innovation. They involve technologies (products) that users can modify relatively easily, and they involve assets that are used repeatedly and frequently.

[The specifics of innovation in the four areas – high performance wind surfing, skate boarding, computer software and games – will be added here.]

In the conventional model of innovation, a manufacturer comes up with a product idea in response to a consumer want or need, develops it into a finished product, and places it on the market. The role of the user of the product (the consumer) is relatively minor in this model, and is largely passive. But the reality is quite different from theory: when we look at what actually takes place, we find that users, both individually and in groups, frequently innovate. Why would users have an advantage in problem solving? Three reasons would seem to present themselves for our consideration. First, asset use provides users with a cheap trial and error process. Second, different users will perceive different solutions to problems. Third, different users will prefer different solutions to problems.

Asset Use Provides a Trial and Error Process:

[This section is in progress.]

Different Users Will Perceive Different Solutions:

[This section is in progress.]

Different Users Will Prefer Different Solutions:

In addition to perceiving different solutions, different individuals will *prefer* different solutions. That is to say, even after multiple solutions have presented themselves, different users will choose differently among them *even if* they have the cognitive capacity to fully perceive the costs and benefits of each one. If the individual is choosing from among his own set of solutions, then his preferences will impact which solution is pursued and disclosed. This can obviously impact the solution set that is revealed to the public. But even if there is a situation in which multiple innovators are revealing multiple solutions, the preferences of innovators can impact the direction and the pace of innovation. [Examples will be inserted here.]

We will now turn to look at a famous example in which innovator preferences had an impact on the direction and pace of intentional innovation: the Wright brothers. Of course, the Wright brothers were not users in the sense that is used here (they were innovating primarily for manufacture, rather than just for their own use). But their example is useful because it is meticulously documented, and because it provides persuasive evidence of both the pervasiveness of the phenomenon and its relative impact. And look at it this way: if preferences - so polluted by the emotion, personality and whimsy of the individual - impact the direction of innovation

among manufacturers and serious inventors, then they will impact the direction of innovation among user-innovators to an *even greater degree*. Another advantage of the example is that innovation in the aeronautics field, particularly in its early days, was particularly dynamic, with many individuals and groups contributing, interacting and competing. This makes the role of choosing among available alternatives (i.e., expressing preferences) particularly apparent. Innovators regularly adopted and built upon the advances of others (even when patents sometimes got in the way), so the impact of preferences can be seen when some otherwise savvy innovators insisted on pursuing or adopting suboptimal solutions.[62]

The Wrights were pioneers in the field of aviation, being the first to achieve heavy-than-air manned flight. They set aviation records, and held the most important patent in the field of aeronautics: the wing-warping patent. But if we look closer at the path of their innovative activity, an unexpected feature emerges. Many times, even after the Wright brothers were made aware of the superior solutions of other innovators, they insisted on blocking these developments (to the extent that they could) and on pursuing their own preferred solutions (and trying to get others to adopt them). Let's take aircraft wheels as an example. All of the Wright brothers' planes were originally designed to take off from a track. Their craft had sled runners, and a weight would be attached to the craft via a complicated arrangement of ropes and pulleys. When the weight was dropped from a derrick, the system would give the vehicle enough thrust to become airborne. It was Glenn Curtiss, another aviation inventor, who first made a plane with wheels, so that a track and a launching apparatus were no longer needed. But even after the Wrights became aware of this innovation and its steady adoption[63] they resisted it: "Orville griped… that pneumatic wheels did not seem like a 'satisfactory' thing to include on a flying machine. 'Personally,' he wrote, 'I think the flying machines of the future will start from tracks, or from [a] special apparatus.'"[64] A similar scenario played out with the placement of propellers: the commonly used placement of propellers behind the wings caused aircraft to stall in flight,[65] resulting in the deaths of a dozen Army pilots. Orville steadfastly refused to change his design to the safer placement in front of the wing[66], because he felt that this would obstruct the view of scouting planes. Furthermore, he was convinced that the stalls were the result of pilot error, and could be prevented if they kept an eye on their angle of attack. He built an 'incidence indicator' to alert pilots to impending stalls, but pilots continued to die, and Orville continued to insist that they needed to pay more attention to the indicator. In the end, the Army outlawed all pusher planes and there were no more pilot deaths due to in-air stalls.[67] But no tractor planes were produced by the Wright Company until after Orville left in 1915.

These episodes were not the result of cognitive difficulties: Orville Wright could clearly perceive the innovations (and their advantages) after they were revealed, but he still preferred his own solutions (even though there is virtually universal agreement that many of the Wright solutions were suboptimal). Nor did these present the sort of 'switching costs' difficulty that might make a manufacturer reluctant to adopt a superior solution because of the cost of

---

[62] A simple approach to suboptimality is adopted. It is defined with reference to things such as planes repeatedly falling out of the air, planes not being able to get off the ground, planes not being able to get back on the ground in less than 3 pieces (an allowance is made for at least one piece breaking off at some point), etc..

[63] They refused to take part in a competition for the Scientific American trophy in 1907 because entered aircraft had to take off from wheels. They did not abandon tracks until [].

[64] Shulman at 128.

[65] I'll insert the explanation here.

[66] This innovation was developed by Glenn Martin of California.

[67] W&O at 388.

retooling: these were the early days when each machine was crafted by hand, and revisions and improvements were constantly made as each new machine and model was produced. These responses can be attributed to a variety of factors. In some cases, Orville was simply determined not to adopt an innovation because it was created by a rival (such was the case with their substandard pontoons, for example). In other cases, he simply weighted certain considerations differently than others (as in the case of the pusher plane). And in other cases, he might have simply been emotionally attached to his own solution. This is not unusual: innovators will often go out of their way to implement their preferred solution. In some cases, this is a virtue: blind persistence in the face of contrary evidence is all that carries many innovators until the day of their vindication. In fact, oftentimes innovations and scientific advances result more from the relentless pursuit of idiosyncratic beliefs and preferences in the face of all available evidence, than from the pursuit of the logical.[68] In other cases, such persistence can slow the pace of innovation, and in the case of the pusher plane, even cause great tragedy. The remedy, of course, is to have multiple innovators finding and promoting alternative solutions. There is one final irony to the story of the Wright brothers: that they were not the ones to put wheels on an aircraft. Why might we have expected this? Because their second vocation, in which they were extremely successful, was as bicycle mechanics. They built, repaired and sold bicycles in Dayton for X years before moving into t Just as we cannot predict how an individual's perception will shape the innovative activity, so we cannot predict what an individual's preferences will be, or what impact they will have on the direction of innovation.

This leads us to the first assertion about user innovation:

> Users of assets will have an advantage in problem solving because: 1) asset use provides a low-cost, repeated, trial and error process; 2) differential cognition across users will lead users to perceive different sets of possible solutions to problems; and 3) differential preferences across users will motivate users to choose different approaches to solving problems. All of these characteristics increase the likelihood of locating a successful solution.

Problem Finding

[This section is in progress.]

This leads us to the second assertion about user innovation:

> Users of assets will have an advantage in problem finding because: 1) asset use will tend to reveal problems; and 2) differential preferences will create different patterns of asset use and so reveal different sets of problems.

The 'Mere' Right to Rent/Resell: Is it Non-innovative?

[68] Insert anecdotes.

[This section is in progress. The evolution of the VCR is complex, with many players trying many different strategies at various times. In general however, the large studios resisted rental early on, wanting to sell tapes for home viewing to the consumer (not allowing rental at all), and later wanting to lease (not sell) tapes to the video retailers for rental.] When unrestricted use of an asset allows users to find new problems and solve old ones, it is easy to see that there is good reason to disallow use restrictions on assets. But what about those cases which appear to be non-innovative? What about the 'mere' right to sell or temporarily transfer the asset without the permission of the original owner (producer)? And what about the right to capture all the profits from these various types of transfers – i.e., what about the law's discomfort with resale royalties? These cases would appear to impact only the distribution of wealth (the way the pie is divided), and not the level of wealth (the size of the pie). Of course, we have already noted one answer to this question: because we do not know beforehand which users will innovate and which ones will not, the law needs a blanket rule that applies to both sets of cases. However, in this section I wish to go beyond this observation and take a look at how the law's prohibition of resale restrictions and resale royalties actually helped to establish a new business model, and even change an industry. To do this, we will take a well-known, and lucrative, example: home video.

Most of us have rented a movie, in some format or other, at some point in our lives. For most of us, our first experience with rental began with the VCR machine and VHS cassettes. Later, we moved on to DVD's, and now some early adopters have already made the transition to Blu-Ray discs. We rent movies from large rental chains, from corner stores, from coin machines and via post. We are a society awash in movie rental. The development of the movie rental business seems inevitable in hindsight, but its birth was marked by controversy and resistance. Who offered the most resistance? Those who now benefit from it the most: the Hollywood movie studios. It seems hard to believe now, but in the early days of the VCR, the major movie studios tried very hard to prevent the establishment of video rental. There were many reasons for this. Many Hollywood executives were simply of the opinion that "rental wouldn't work"[69]. But they were also worried that rental would facilitate piracy, and that rental would become a cheap and convenient alternative to a visit to the movie theater – thus resulting in a cannibalization of box office revenues. They were willing to sell high-priced VHS cassettes to the serious movie buff, but they didn't want these tapes to be rented. Others were willing to allow rental, but only if they got a cut of each and every transaction. These concerns gave rise to various strategies on the part of different studios. Some partnered with consumer electronic firms to develop early formats that incorporated anti-piracy features and had the potential to maximize revenues. Cartrivision, for example, used pre-recorded video cassettes that could not be rewound. This not only stymied piracy, but made consumers pay for each and every viewing. Others promoted non-recordable formats, such as laser discs. Others put 'no-rent' clauses in contracts that accompanied the transfer of tapes to distributors; while others sought to develop exclusive partnerships and 'lease-only' deals with the rental outlets themselves. It is easy to see why the forays into anti-piracy formats were spectacular failures: they simply offered too little functionality and value when compared to the VHS format. But why did the contractual/transactional strategies also fail? It was here that property law played a starring role. The Copyright Act's first sale doctrine prevents a copyright owner from controlling the use and distribution of a copy of a work after it has been sold. This means that you can watch a movie, or read a book, as many times as you like without paying the copyright owner more money (or asking for her permission). Similarly, you can sell,

[69] VVV at 113.

or lend, or rent, your VHS tape or book to someone else if you so choose. Thus, anyone wishing to set up a movie rental shop could do so without the permission of, and without continuing payments to, the Hollywood studios. And this is exactly what video retailers did. Of course, the *distributors* could be held to the 'no-rent' contracts they had signed, but third-parties (the retailers) could not - and they were the ones doing the actual renting. Leasing agreements were another matter: they could be enforced because they were made directly between the retailer and the studio. But the momentum that first sale provided caused retailers to resent these leasing agreements, and they were abandoned after only a few years. The only thing left for Hollywood to do was to charge more for the original sale of the cassette, and although this initially kept prices on some popular movies higher than they might have otherwise been, it enabled film independents to compete on price where they could not compete on content. Independents were only too happy to supply retailers with restriction-free, reasonably-priced cassettes. This new revenue source enabled independents to expand production and become significant players in the financially risky world of film production.

The innovation story here may not be a traditional one, but it is an important one. Video retailers were not creating a new product in the usual sense. Instead, their innovative contribution resided in the development of a new business model, and in the satisfaction of consumer wants, back when no one knew whether video rental would work (and when many Hollywood players were convinced that it wouldn't). Hollywood would most likely have gotten us there eventually, it is true, but the first sale doctrine got us there much faster. And video rental had other downstream effects which no one could have predicted. [Insert summary of how VHS revenues financed the rise and fall of the independents.]

And in case there are any lingering doubts in readers' minds about the role that the first sale doctrine played in creating the video rental market, we can contrast movie rental with music rental. Most of us have rented a VHS tape or a DVD, but very few of us have rented an audio cassette tape or CD. Why? It's not because some savvy entrepreneur did not come up with the idea. In Japan, a record rental shop was first opened in 1980. And it's not because consumers did not respond to the idea. Within the year there were more than 34 shops in Japan; by 1984 there were more than 1,600. So why did record rental not evolve in the robust manner of movie rental? The reason is that the first sale right with respect to audio CD's was modified after intense lobbying by the Recording Industry Association of America (RIAA). In 1984, RIAA was successful in convincing Congress to pass the Record Rental Amendment Act which made it illegal to rent (though not to resell) records. At that time, there were only a handful ("three or four"[70]) of record rental stores in the U.S., so although there was little practical impact in this country, RIAA used the evidence from Japan (they asserted that record rental had caused a 30% decrease in record sales) to convince Congress that record rental had to be prevented. The movie studios also took part in these hearings, but they were unable to produce any evidence that movie rentals were harming their bottom lines. More importantly, by 1984 there were [] video rental outlets in the U.S., so members of Congress would risk the ire of their constituents if they gave Hollywood control over movie rental at this late stage in the game. Thus, the final tactic in Hollywood's anti-rental arsenal - repeal of the first sale doctrine - failed, and the video rental market grew at an astonishing pace throughout the 80's and 90's.

The example of home video is germane because it illustrates how complex changes in industry structure (worldwide Hollywood dominance, industry consolidation and the blockbuster model of movie production) can result, in part, from simple antecedents in property law (first

[70] VVV at 115.

sale). As an interesting thought experiment, we might think about how record rental might have catalyzed a different trajectory of the recording industry, just as it did for the movie industry. Of course, it is impossible to say what might have happened in this hypothetical alternate universe, but it remains an interesting question nonetheless. After all, no one could have predicted how home video fundamentally changed the film industry – and indeed, no one could have predicted the precise sequence of events that resulted in those changes. Another nice thing about this example is that it provides an illustration of how a user can more easily move from being a passive consumer, to offering a desired service or product to consumers, simply by building on the legal institution of sale. No level of technical sophistication is necessarily required for consumers to become actively involved in the new markets they help to create. (Recall that users who innovate via product modification and idea generation sometimes become manufacturers themselves, or sell their innovations to manufacturers.) Of course, the most ironic thing about this example is that it illustrates how the first sale doctrine fostered the growth of an industry built on *rental*. Nothing about that fact mitigates its usefulness as an example. The *numerus clausus* does not operate to get rid of governance rules in the case of temporary transfer, only upon permanent transfer. Furthermore, even in the case of temporary transfer, full use tends to travel with possession for that limited period of time.[71] That is to say, even in the case of rental, consumers can watch (use) the movie as many times as they like for the period of time in which they possess a copy of it. The significance of this is examined in more detail in Part 4.

Accidental Innovation

Although we often associate innovative activity with the type of directed search that characterizes intentional innovation, innovation can also be unintentional, or accidental. Indeed, accident has played a pivotal role in many important innovations. There are many examples: Velcro, dynamite, Teflon, penicillin, Rayon, anesthesia, Ivory Soap, Corn Flakes and Scotchgard - are just a few of the more well-known products which mistakes, oversights and luck have brought to us. Accident enters the innovative process in many different ways. Sometimes a mistake is made in executing an experiment; sometimes a researcher is looking for a particular product and finds something completely different; sometimes an 'accident' occurs in the classic sense – spills, trips, cracks and leaks. Let's take the chemical industry as an example, which has been greatly aided by the clumsy and the careless. Indeed, if chemists didn't drop things, or forget to clean things, the chemical industry might have ground to a halt long ago. Of course this is a deliberately facetious exaggeration, but there is a grain of truth in it: broken thermometers, spilled solvents, leaky equipment – all these have contributed to the some of the most useful innovations of our time. A commercially feasible method of manufacturing synthetic indigo was developed when a broken thermometer spilled mercury into a reaction vessel. This allowed Britain to enter the indigo market, and destroyed India's dominance in the market (which up till then had relied on natural indigo from plants). The chemical used to Scotchgard fabric was developed after a chemical was spilled on a tennis shoe, and that area of the shoe resisted soiling. Polyethylene (used in a wide variety of products from garbage bags, to product packaging, to underwater cable insulation) was synthesized by accident when leaky equipment allowed just the right amount of oxygen to enter a sealed, high-pressure reaction. Even errant lab members have

[71] Of course, copyright law reserves certain uses to the IP owner, e.g., public performance of the movie. Private performances however, are not so reserved.

played a role. Nylon, one of DuPont's most profitable products, resulted from the combination of an abandoned experiment, an unsuccessful experiment, and most importantly - some fooling around by lab members.[72] A member of DuPont's basic chemical research lab, Julian Hill, had noticed that a polyester polymer stuck to his glass rod when he dipped it in the material, and formed silky fibers when he drew it out of the flask. When the head of the lab, Wallace H. Carothers, went to town, Hill and his lab mates decided to run down the hall to see how far the material would stretch between them. When they did this, not only did they notice that the material did indeed stretch to an entertainingly impressive length – they also observed that the strands seemed to grow stronger (!) and more 'fiber-like' as they were stretched. They realized that the stretching was causing the molecules in the polymer to line up and form bonds, and that such a process might be used to produce a useful synthetic silk. Was the material used in these lab high-jinks Nylon? No. It was a polyester that was not useful in textile manufacturing.[73] But their observation of the 'strengthening by stretching' process (called the cold-drawing process) led them to try it on a previously abandoned material, Nylon - a synthetic analog to silk. This led to the wildly successful material used in parachutes and (most famously) women's hosiery.[74]

The role of accident in scientific discovery and industrial innovation has been well documented. Unfortunately, the role that it plays in user innovation more specifically has not been so carefully chronicled. Often, even when accident does play a role in user innovation, it is not even identified as such. Let's take another well-known product as an example - baking soda. Most of us have at least one box of this in the refrigerator - we use it to absorb odors. But this was not always the case. From 1846 up until the late 1960's or so, baking soda was used to bake, not to make our refrigerators smell better. The primary brand, then as now, was Arm & Hammer, and it was one of Church & Dwight Company's flagship products. How did the predominant use of baking soda change from baking ingredient to refrigerator deodorant? A combination of accident, product obsolescence and user innovation. You see, once the 1970's rolled around, the sales of baking soda were declining because many working women no longer had time to bake 'from scratch'. They switched to baking mix, or began to buy pre-made cakes and cookies. So, what happened to that now almost useless box of baking soda? It was taken out of the cupboard and put in the refrigerator. Perhaps hoping that they would occasionally have time to bake, perhaps not wanting to throw an unspoiled product away, the women put the unused product where it 'would keep'. And then something unusual happened – people began to notice that it absorbed odors, and as 'word of mouth' spread about this new use (likely in part via friends and family, in part via tips in women's magazines) more people began to put it in their refrigerator. Employees at Dwight & Church heard of this new use, and did some research. What they found confirmed what users had already discovered – their product did indeed absorb odors. The company responded with an ad campaign in 1972 that marketed Arm & Hammer Baking Soda as a refrigerator deodorant, and in a few short months, "more than half of American refrigerators"[75] had an open box of baking soda. This might seem like a rather simple example of user innovation – it literally involved nothing more than an unintended discovery of a new use of a product. But it illustrates 4 features of interest to us: 1) the innovation was accidental; 2) the innovation was

[72] Chronicled in Serendipity, Chapter 25.
[73] Later other useful polyesters were found, and were used in a variety of products, including textiles.
[74] Here, it was the process (cold-drawing), rather than the material (Nylon) that was patented. The material was useless without the process, and the process was later discovered to be useful on a variety of materials.
[75] Sawyer, 470 and C&D history.

user driven; 3) the innovation was successfully adopted by a manufacturer; and 4) the innovation was economically important [insert sales figures here].

Lack of documentation is only one of the difficulties in assessing the role of accident in user innovation. Another is identification. Quite simply, the line between intentional and accidental innovation is not always a bright one, particularly when user innovation is involved. This is best illustrated by the cases in which asset use naturally reveals a new problem or solution (which seems more akin to accident), which is then followed by an active search for a solution to the found problem, or active refinement of the found solution (which seems more like intentional innovation). Is the final innovation in such cases a result of intention or accident? It would seem to be a bit of both. If accidental and intentional innovation often occur together, then it may not be strictly necessary to document specific instances of accidental user innovation – we can assume that at least some cases of intentional user innovation are triggered by accident. In other words, *use* can reasonably be considered to be a *good generator of accidents.* Accident can enter the process of innovation in various ways. Sometimes, an individual is not trying to innovate, but an accident reveals something new. At other times, an individual might be looking for something and find something else instead. Realizing the value of accident, some individuals even seek to intentionally inject randomness into their search. Finally, an individual might make unexpected connections between unrelated things or concepts.[76]

How Can We Account For The Role Of Accident In The Innovative Process?

In many ways, the idea of accident may seem antithetical to the idea of innovation. By their nature, accidents represent unexpected occurrences, so how can we 'plan' a program of innovation around them? Furthermore, accidents are most often unwanted, as they represent divergences from what we expect to happen. We spend time and money preparing for things to turn out as we reasonably expect, for events to unfold in a manner that, while not completely predictable, is reasonably intelligible in terms of past events. Accidents represent a break in this orderly unfolding, and thus can be quite costly. At the very least, it seems inefficient, and at worst, foolhardy to pursue innovation via the vehicle of chance, luck, or accident. But there is also another way to look at the relationship between accident and innovation. Instead of polar opposites, we can view one (accident) as a normal and frequent part of the other (innovation). Indeed, many scientists, from Ernst Mach to Salvador Luria, have expressed a conviction that accident plays a significant role in innovation. Consider this from the Nobel Laureate Sir Alan Hodgkin: "I believe that the record of [my] published papers conveys an impression of directness and planning [in my work] which does not at all coincide with the actual sequence of events. … [O]ver a long period I have developed a feeling of guilt about suppressing the part which chance and good fortune played in what now seems to be a rather logical development."[77] A useful way to formalize these intuitions can be found in the work of Donald Campbell. He theorized that all creative thought (and all knowledge processes) is a combination of two steps: 1) blind variation; and 2) selective retention. In part, this conception of the process of creative thought is driven by the inherent difficulty of accounting for the psychological process that produces knowledge that is truly new. We need some mechanism - 'blind variation' – which enables an organism to generate knowledge which is discontinuous with the previous stock of knowledge. This is the

[76] These are adopted from Austin & Devin.
[77] Pursuit of Nature: Informal Essay on the History of Physiology, Cambridge, at 1.

only way we can account for "repeated 'breakouts' from the limits of available wisdom."[78] It is easy to see why this is a useful framework for thinking about creativity or innovation, which also needs a way to account for the development of that which is completely new.[79] Thus, although Campbell was initially concerned with the blind variation that occurs in the thought processes of any one individual,[80] his ideas can also be applied to variations that we encounter in the external environment. To understand the difference between the two, it might help to note the following: In the case of creative thought, the thoughts of the organism are independent of the environment (thoughts have no antecedents in, and are not in response to, the environment). In the case of accidental creativity or innovation, events in the environment are independent of the thoughts of the organism (events have no antecedents in, and are not in response to, thought). Accidental creativity or innovation includes cases involving creative thought, but *also* includes those cases in which events in the environment are independent of the thoughts of the organism. This gives us a complete theory of accidental creativity or innovation – one which accounts for new ideas (creative thought) and new applications (creativity or innovation). This extension of Campbell's work is consistent with scholarship on accidental creativity and innovation, which has built on the insights of his model.[81] Perhaps more interestingly, it also corresponds to the informal theories of innovation put forward by many scientists, such as W. B. Cannon: "Obviously, a chance discovery not only involves the phenomenon to be observed, but also the appreciative and intelligent observer."[82] All of these share the two-part structure in which randomness first produces multiple variations, and an observer then chooses among them. The ability to choose intelligently among the variations (selective retention) has been referred to in various ways: 'prepared mind'[83], 'intelligent observer' and 'sagacity' are just a few of them. The term 'sagacity' will be used here.

The generation of a large number of potential thoughts, ideas, variations or events is only the first step involved in accidental innovation. Does the second step, the requirement of selective retention, imply that only sophisticated individuals (such as scientists or industrial researchers) are capable of building on the opportunity provided by accident? We have already seen the sagacity of users displayed in the pursuit of intentional innovation. There is no reason to think that this sagacity would not also enable users to successfully recognize the significance of fortuitous accidents. Indeed, given the sheer number and diversity of users, they would seem as a group, to have a distinct advantage in terms of their capacity to generate both novel thoughts and novel external circumstances. And only a relatively small number of them need to successfully recognize and pursue those situations of novelty for society to see useful results. This accords with the view put forward by Campbell: "Indeed, looking at large populations of thinkers, the principles make it likely that many important contributions will come from the relatively untalented, undiligent, and uneducated, even though on an average contribution per capita basis, they will contribute much less."[84]

---

[78] Campbell at 380.

[79] Discontinuity (without antecedent) does not imply that the innovation be 'radical' in terms of impact.

[80] Though Campbell did mention the parallels between his theory of creative thought and Mach's thinking on innovation.

[81] Actually, Campbell's blind variation-selective retention model has been applied to a multitude of evolutionary processes: organizational science, cultural evolution, [].

[82] Cannon, "The Role of Chance in Discovery" at 207.

[83] The famous quote from Pasteur is often recorded as: "Chance favors the prepared mind." It is actually: []

[84] As an aside, this claim is made about Fleming, who helped discover penicillin – that he was a capable experimenter, but not all that great a scientist. Eurekas at 182.

Keep in mind too, that we have already seen that intentional innovations are often produced by non-experts, or by people untrained in specific disciplines. Hence, the 'prepared mind' referred to by Pasteur need not refer to any particular level or type of education or preparation. Indeed, to the extent that this sort of preparation narrows an individual's willingness to look beyond expected outcomes, it might even work to prevent him from recognizing that something unusual could be important. This is not to say that classical training is always, or even often, harmful. Take the DuPont chemists as an example. It was their training in chemistry that enabled them to recognize that stretching the polyester oriented the molecules and led to the formation of bonds. This allowed them to recognize (selectively retain) the importance of the unusual event (the blind variation - which was the strengthening of the material upon stretching). So, I do not mean to suggest that experts will be *less* sagacious than the non-expert, but only that we need to have a definition of sagacity that goes beyond classical training, and embraces the full range of idiosyncratic perceptions and thoughts of the 'common man'. Of course, the field to which the innovation contributes will also help determine the relative contributions of experts and non-experts, a point which is also noted by Campbell.[85] (We might expect few contributions to theoretical physics by the uneducated, many more contributions to product modifications and computer software, for example.) However, it is important to note that the sophistication of the innovator does not necessarily correspond to the importance of the innovation. Take quinine as an example. It was discovered sometime before 1630, [86] and remains *the best* modern cure for malaria – which is still one of the most lethal diseases worldwide.[87] Although the exact circumstances of its discovery are uncertain, it is believed that it was discovered when a feverish and disoriented South American Indian accidentally drank water that was infused with the bark of the quina-quina tree. The bark was thought to be poisonous, but the Indian not only did not die – his fever was cured. From then on, quinine was used to prevent and treat malaria. This story is difficult to verify in its exact details, but even if quinine was not discovered in this way (though someone had to be the first to drink it), there are many important and effective natural compounds that were unintentionally discovered by tribal societies. Human beings have always had the sagacity to learn from accident, so there is every reason to suppose that a large and diverse group of modern users would display an even greater ability in this regard.

Is Accidental Innovation Merely A Curiosity, Or Is It Something Worth Fostering?

The anecdotal evidence offered above makes it clear that at least some scientists are convinced that accident is the handmaiden of innovation. And although these men were speaking about the role that accident plays in scientific *discovery*, their comments apply to an even greater degree to the development of innovative applications of knowledge (*inventions*). Why? Because accident merely speeds up the pace of scientific discovery, but it literally pulls inventions from the thin air. Someone would have eventually discovered X-rays, for example, even though Röntgen discovered them much faster thanks to an accident. But would someone have inevitably created Velcro, whose invention stemmed from George deMestral's frustration with removing cockleburs from his coat? How about floating soap (Ivory Soap), which was created when a

[85] Ibid.
[86] This is the first documented usage. See Serendipity, Chapter 3.
[87] There are several synthetic versions of quinine, but new strains of the malaria parasite have developed, and some of them are resistant to them. No strains have yet been found to be resistant to natural quinine.

worker mistakenly left a stirrer on during lunch? The air-filled batch was almost discarded (generally, we want soap to be more dense, not less), but was sold because of corporate parsimony. Only when consumers wrote in requesting more of that 'floating soap' was a new product, 'Ivory Soap,' born. The point here is this: as we move from the discovery of what exists but is unknown, to the invention of what did not previously exist, accident will play a more pivotal role. Of course, even the discovery/invention and known/unknown distinctions do not provide airtight guides as to whether accident merely speeds things up, or gives us what otherwise would not exist. A material or substance might exist and be known, for example, but an important use or application, might go undiscovered (un-invented?) but for a fortuitous accident.

Scientists and inventors are not the only ones who have extolled the virtues of accident. Artists have also expressed a keen appreciation for the role of the unexpected in the creative process. Some artists seek to intentionally inject randomness into their work, for example, by resisting the urge to plan or structure their work process. Other artists have stated that they do not 'create' at all – that they merely choose among the random thoughts, melodies, etc., that enter their minds.[88] Some businesses and their analysts are also beginning to recognize and actively promote the role of accident in business innovation. Business school scholars are also beginning the task of analyzing and formalizing accident's potential for generating innovations and increasing profits. Indeed, some scholars have suggested that there may even be a positive relationship between the extent to which we allow accident to insert itself into the innovative process, and the *degree* of innovation that results. Robert Austin and Lee Devin, both of Harvard Business School, hypothesize that truly breakthrough innovations generally result from searches that are unplanned and flexible (i.e., highly inefficient), because accidents reveal what planning cannot. They note that the more traditional model of business innovation, which involves tighter controls and more careful planning, yields innovations that are often more incremental in character. There is some evidence to support their hypothesis. The criticism that too much planning (too little accident) has worked against innovation has been aimed at the pharmaceutical industry, for example. Some commentators have noted that the use of computers in the search for new drugs has hindered, rather than helped, in the search for new drugs. Automation of structural analysis and materials testing has produced an overall greater number of drugs, but fewer breakthrough drugs – i.e., fewer drugs which really make money. How can this possibly be so? Computers can propose new molecular structures for possible drugs, and can test the properties of new chemical compounds far faster than human beings. So what is happening? Quite simply, new drugs are often discovered by accident. This happens at every stage of drug development – synthesis, testing, and clinical trials. [Insert examples.] Human beings are far slower than computers – they process fewer data per second - but they have other advantages. One of the most important is the ability to notice and analyze things outside a given set of parameters. They can also make connections that, before they are made, simply do not make sense. In other words, human beings can make mistakes, and mistakes can give rise to something new and completely unexpected. Making mistakes, seeing things 'wrong' – can literally be a virtue.

Finally, in assessing the importance of accident in the innovative process, we should consider the fact that some mathematicians have asserted that accident is actually responsible for most of the major advances in human history. [Insert some examples.] Although these advances can be understood in terms of a series of connected events after they occur, they could never

[88] Campbell at 387.

have been predicted in advance. This means that they could not have been planned for, or made to occur – they had to occur by an accidental sequence of events. If this is true, then accidental innovation may be far more important than we think - it may even be more important than planned innovation. We have seen artists and scientists, scholars and businessmen, experience and theory, line up in support of the significance and irreplaceable nature of accident. Taken as a whole, the variety and venerability of the foregoing evidence makes it reasonable to assert that accidental innovation is important enough to cultivate.

How Do We Foster Accidental Innovation Among Users?

Given the nature of accidental innovation, there is no way for us to 'arrange' for specific accidents to occur. Nonetheless, we *can* create conditions under which accidental innovation is more likely to happen. Fostering accidental innovation means creating conditions under which: 1) more accidents are likely to occur 2) to a wide variety and a large number of people. Such conditions satisfy both requirements of Campbell's model. A larger number of accidents means more blind variation (more potential innovative opportunities from which to choose). A greater number of people (and more variation amongst them) increases the number of cases in which the observer of a particular accident has the sagacity to recognize its importance (selective retention). It is easy to see how the law's technique of using tangibility as a rough proxy for human capital might help us to create just these types of conditions. When an asset permanently changes hands, the law prefers to give the possessor (owner) of that asset full and free use of that asset, even when the transferor attempts to make that possessor a mere licensee who is subject to use restrictions. Use would seem to be a major source of accidents, so disallowing use restrictions on assets allows more accidents to occur. The over-inclusiveness of the legal rule is also crucial: we do not know which individuals will have the cognitive capacity to identify useful accidents, and the motivation to build on them. But the law applies rules of exclusion (which give the possessor free and full use of the asset) no matter what activities users engage in – intentional innovation, accidental innovation, failed innovation, or no innovation. There is no *ex ante* inquiry into what any particular user might do, or *ex post* inquiry into what any particular user has actually done. This increases both the number, and the variety, of people to whom accidents might occur, thereby increasing the chance that accidental innovation will take place. Thus, efficient breach in property law, which appears inefficient because it applies across the board, is actually more efficient (value enhancing) in the long run – precisely because it applies across the board. Indeed, allowing users to engage in unrestricted 'creative play' with the assets that they possess (own) will increase *both* accidental and intentional innovation.

The concept of 'creative play' might seem too frivolous to be of assistance in the great and ponderous march of progress and innovation, but some similar suggestions have actually been aimed at industry. So, for example, Austin & Lee advise businesses to move out of their 'cone of expectations' by, among other things, inducing accidents, playing with unusual combinations, reducing modularity of tasks, and collaborating with individuals of diverse backgrounds and experience. Of course, reliance on increased flexibility and decreased planning entails its own set of costs. In the case of business innovation, the cost of a more flexible approach is greater uncertainty and potential inefficiency. In the case of user innovation, the cost is the disruption of the carefully designed business model (or the pattern of investment) of the seller. Settled expectations are the province of property law, it is true, but they are the enemy of

accident. Actually, it is more accurate to say that they are the enemy of innovation. And to the extent that accident is the ally of innovation, property law should be structured to allow accidents to occur, and for people to take advantage of them. That is exactly what the efficient breach structure of property law, via its preference for rules of exclusion, does.

There is a related concern in all of this, which actually applies to both accidental and intentional innovation. If, over time, users get accustomed to *not* interacting freely with assets, this can have a deleterious impact on the level of innovation (no matter what the law happens to be). How might users become less likely to use assets in ways not contemplated or countenanced by the seller of those assets? It could happen in a variety of ways: 1) if users are afraid to breach use restrictions of which they have notice because of the threat of getting sued; 2) if users find it difficult to breach use restrictions; or 3) if society adopts a norm of compliance with use restrictions, and users internalize that norm. The first of these scenarios is already occurring - the 'chilling effect' of lawsuits has been noted in the case of computer software, for example. [Insert examples here.] The second scenario is also occurring, at least in the case of digital and electronic goods. These goods can be designed with built-in restrictions that limit the range of uses/modifications that consumers can make of the good. The third scenario is more speculative, though at least some people hope that it will someday become a reality. The movie industry, for example, spends an enormous amount of time and money seeking to shape consumer habits.[89] Most of us have seen the 'anti-piracy' ads that run in theatres and at the beginning of DVDs – there is little reason to think that these will alter consumer habits and attitudes (and there is little to concern us in the content of these ads). Of more concern are elaborate education programs aimed specifically at children, designed to alter how they think about interacting (and how they actually interact) with IP works. [Insert examples here.] There is little that property law can do to respond to the first (litigation) and last (education/advertising) of these scenarios, but some legal responses to the second (technological restrictions), will be discussed in Part 5.

This leads us to the third assertion about user innovation:

> User innovation can be high-tech or low-tech; intentional or accidental; IP-producing or information-producing. Whatever the type, user innovation will be more likely to occur if users are given free and full use of the assets they possess.

Efficient Breach: A Contentious Concept

Now that we have seen how the replacement of governance rules by rules of exclusion can foster user innovation, thus resulting in a form of efficient breach, it is appropriate to take a step back and consider some objections to this conceptual analogy. As a general matter, it may seem unwise to build a theory of one of the core features of property law on the back of a doctrine that is generally considered to be contractual in origin. However, there is an even more

[89] And indeed, with technology comes the possibility of doing this in a hidden way. Technology shapes consumers' perception of what freedoms they should have, by shaping what freedoms they actually do have. More about this in Part 5.

pressing charge to answer on this point, because the status of efficient breach, even within contract law, has been the subject of great criticism and debate. The objections to efficient breach in contract law can be organized under five main headings: 1) efficient breach does not exist; 2) efficient breach should not exist because promise is the basis of contract; 3) 'efficient' breach is not actually efficient; 4) 'efficient breach' should only exist if there is an analogous concept in property law, and there is none; and 5) efficient performance and efficient breach should produce invariant results, and given the choice between them, we should prefer efficient performance (see objection number 2). We will briefly[90] review each of these objections in turn.

1) Does Efficient Breach Exist in Contract Law?

Several scholars have argued that the Holmesian version of contract law is a mirage, more a creature of academic seduction than practical application. That is to say, it exists in the minds of legal theorists and the spilled ink of legal theory, but not in the courts. There is some merit to this view. It is admittedly true that the courts have never embraced efficient breach with the vigor that Holmes envisioned, even when they pay lip service to the principle. That is to say, courts are regularly influenced by the implicit morality of keeping promises when setting damage awards. Even some law & economics scholars have shied away from wholeheartedly defending the willfully breaching party. Despite the fact that damage awards fail to reflect the full austerity of Holmes, neither is it correct to say that efficient breach receives no level of recognition in the courts. That is to say, until courts require specific performance as a general matter, or disgorgement becomes the standard measure of damages, then breach does have a secure place in contract law - in fact as well as in theory. [This section is in progress.]

2) Should Past Promise Trump Present Freedom?

The concept of efficient breach has also been criticized in contract law on the basis that the morality of keeping a promise should be paramount in determining the rights and remedies as between the parties to a contract. The ability to breach is a huge fault line in contract theory, with heavyweight scholars lining up on both sides of an unbridgeable divide. It is difficult to know how to respond to this objection in this work: there is either little one can say about this issue (because, as a matter of policy, you simply don't accept this version of contract), or far too much (because you have a complex set of reasons why you think this version of contract is bad policy). Contract law is replete with doctrines that operate to allow contracting parties to escape their deals: unconscionability, concerns about inequality of bargaining power, duress, mistake, frustration; all of these operate to legally undo what both parties agreed to, but which one party now regrets. Even the concept of agreement provides an escape valve from performance. Given this, it is difficult to know why breach, the most obvious means of escaping performance, should be so singled out for approbation. [This section is in progress.]

3) Is Efficient Breach Really Efficient?

It has also been asserted that 'efficient' breach is not even defensible on its own terms – that 'efficient' breach is not even efficient. There are two major objections to the assertion of

[90] Quite briefly, in fact, since an entire work could be written on this issue alone.

efficiency: 1) that the efficiency calculus does not take the costs of determining expectation damages into account; and 2) that we can never say for certain if breach is really efficient or not. The first point can be easily answered in the case of property law – there is no need to determine damages at all, since the application of the *numerus clausus* involves giving a full property right to the downstream possessor. The second point is more a statement of agnosticism than a real objection: "Admittedly, in some situations the breach (as well as other unlawful appropriation) might enable the wrongdoer to *derive gains exceeding the victim's loss*. But the prospect of some gain does not turn the unlawful appropriation in to a lawful one."[91] If the breach allows a resource to move to its higher-value use, then breach is indeed efficient, whether we consider the actions of the breaching party to be morally reprehensible or not. Furthermore, it is the law which determines whether breach is unlawful or not, so if breach is part of the law of property and contract, then the appropriation is not 'unlawful'. The gains of the breaching party (or of the new property owner) never 'belonged' to the 'victim' in the first place.

4) The Lack of a Property Analogy

The fourth objection that is often raised against the concept of efficient breach in contract law is that there is no analogous concept in property law, which there should be if the concept has any validity.[92] The analogous concept is usually referred to as 'efficient theft' or 'efficient conversion'. While it is true that there is no such thing as efficient theft or efficient conversion in property law,[93] the above discussion of governance and exclusion has shown that there *is* such a thing as efficient breach. In property law (at least in the *numerus clausu*s version of property law), we breach the governance rule, whether that governance rule is considered to be a license, a servitude, or other encumbered property form. Unless and until we make law over in equity's image, then the concept of efficient breach – in property law, *even more than* in contract law – will remain an integral part of the very meaning of the right. In fact, if we could replace our rough bundles of 'property' (rules of exclusion) with an endless web of payments and permissions (governance rules), property as we know it would not exist at all.[94] Of course, if we didn't have the concept of sale, or the *numerus clausus*; if first in time really *was* first in right, then there would indeed be no 'efficient breach' in property law. But that is not what we have, and that is not what we should have.

Viewing the legal version of property as a form of efficient breach might at first seem odd, but even from a state of nature, property operates as a form of efficient breach. Consider a hypothetical country in which all property was owned by the entire community. In order to develop a piece of land, A would first have to get the consent of all members of the community. Instead of viewing this scenario as shared ownership situation, however, we can also view it as a complex web of bilateral agreements: A has an agreement with B not to take from the commons without B's permission; A has an agreement with C not to take from the commons without C's permission, etc. If this community sets up rules under which property can move from the commons to the private sphere without the consent/compensation of each member, property

[91] Friedmann at 13.
[92] See, e.g., Friedmann, Epstein. We heartily agree with the assertion that the concept of efficient breach should be valid in property law if it is valid in contract law. Of course, the answer here is that the concept is indeed valid in both.
[93] Actually, adverse possession might qualify as some form of these.
[94] More on this in Part 4.

essentially operates as a form of efficient breach. Instead of abiding by his many agreements with B, C, D and so on, A can now simply vest himself with ownership (as per the rules specified by the community) and go about unilaterally determining the best use of the property. The notion that private property lowers transaction costs is a familiar one to law and economics scholars,[95] and it is relatively easy to see the cost-saving advantages of this type of efficient breach when moving from communal to private ownership. But once we move from communal ownership to private ownership, there is no longer an issue of excessive fragmentation of rights (there are only two parties now, the transferor and the transferee), the rationale seems to break down. There seems to be little harm (and some good) in allowing A to create customized servitudes, for example, because they allow him to efficiently coordinate the long term development of land parcels. A will be much more likely to invest in improvements in his own land if he can be sure that his neighbor cannot destroy its value by, let's say, building a 30-story eyesore in a residential area. Servitudes can thus act as a kind of private zoning. Similarly, servitudes on chattels can entice producers into the market because they protect investment in new business models. If I produce both toner cartridges and printers, selling my printers at a loss in order to sell more cartridges, then I need to have some way of ensuring that consumers need a steady supply of cartridges. One way to do that is to through creative use of servitudes: a notice might be placed on every cartridge, stating that the cartridges cannot be reused (i.e., refilled). Such scenarios are appealing in their orderliness and their defense of well-established plans. But when we shift the focus from what restrictions a former owner might want enforced, to what a present owner might do if given an unencumbered asset, then we can see that even in the bilateral case, user innovation can result in a form of efficient breach.

5) Why 'Efficient Breach'? Why Not 'Efficient Performance'?

The observation that B's cost of searching out new uses of an asset is lower if he owns all rights in the asset, does not in itself explain why it matters that B, rather than A, owns the right to permit a new use. Why won't A and B make a deal if B's new use makes everyone better off? The Coase theorem would suggest that unless transaction costs are prohibitively high, parties will bargain their way to an acceptable rearrangement of rights, and so it does not matter how the rights are initially allocated (this is the invariance thesis[96]). Although this assertion is most often made vis-à-vis the initial assignment of property rights, it can equally be made about the issue of breach in contract law. Recognizing this, some commentators have asked the question of why contract law, rather than allowing 'efficient breach', does not instead require 'efficient performance'. This would give promisees an enforceable right in performance, which could then be bought out by the potentially breaching party. After all, if the potentially breaching party is rational enough to see the benefits of breach, why wouldn't the promisee see the benefits as well, and be satisfied with a share in the profits from that breach? The assignment of rights should generate invariant results in contract and property transactions alike, and there should be no difference in the efficiency of granting a right to specific performance (of a contract *or* a

[95] Posner, Epstein, Past and Future at 670.

[96] This is actually the 'soft' version of the invariance thesis. The strict version states that the distribution of wealth, as well as production, will be invariant under any initial assignment of rights. See, e.g., Coase, Notes on the Problem of Social Cost, at 171: "a change in the liability rule will not lead to any alteration in the distribution of wealth." No comment on that version of the thesis is ventured here. See, Stigler, Cheung for additional commentary.

servitude), or allowing breach (although the distribution of wealth certainly differs in these two scenarios).

The invariance issue has been hotly debated by both legal scholars and economists, and commentators have thus far identified two principle sources of transaction costs which may cause the thesis to break down: valuation and detection. Much has been written about how assets that are intrinsically difficult to value - marital assets or intellectual property, for example - can make it impossible for parties to reach agreements over them. In the case of marital assets, sentiment (and resentment) may generate large variances in parties' estimations of asset value. In the intellectual property case, it may be simply difficult to estimate how much value a particular IP input adds to the value of a finished product. Such difficulties can give rise to strategic behavior which causes bargaining break down. Detection is a quite different matter: it may be difficult, in certain types of cases, for parties to even know that their innovative activity has created an externality that has to be bargained over. Once again, this type of detection problem is generally much greater for intellectual property assets than it is for tangible assets. There are actually two types of detection problems in the IP context: the first involves a problem of asset revelation; the second involves a problem of asset duplication. The first problem is well-known as Arrow's information paradox: if I create an information-intensive asset, I need to reveal it to you so that you can know what I am offering to sell to you. But once I have revealed it to you, you have the information and you no longer need to buy it from me. IP rights facilitate disclosure by granting a property right to demand payment for use (or stop the use) of an asset, whether or not a deal is struck. The second type of problem is solved in very different ways by different IP regimes. A cloistered innovator can slave away in his garret and unintentionally create a work that is quite similar to one already in existence. If that innovator is an author, copyright law vests his duplicate creation with its own property right – thus doing away with the detection problem completely. If that innovator is an inventor, on the other hand, and he has violated a pre-existing patent, then he is liable for infringement whether or not he knew he was 'using' someone else's invention. Chattels can be considered another example of a good which has difficult detection problems. Even if I want to restrict certain uses of my product, it may be difficult for me to know when consumers are making such uses. Of course, even in the case of tangible assets, detection is becoming somewhat easier via clever use of technology, which enables better tracking and control of specific uses.

Although the hurdles generated by asset valuation and infringement detection are doubtless important in a significant number of transactions, they are insufficient, by themselves, to account for the presence of efficient breach in property law. There are two reasons for this. First, they only apply to very specialized types of assets and transactions, and therefore do not cover the broad range of assets and transactions regulated by the *numerus clausus*. This first deficiency will be addressed in Part 4, where problems arising from disputes about valuation and strategic behavior will be broadened to include the complex endowment effects that arise in relation to all sorts of assets - tangible and intangible, moveable and immovable. Problems related to detection, because they are essentially problems of information transfer (and are thus taken care of by notice), will also be discussed in Part 4. What we are concerned with at this point in the argument, however, is not a problem of valuation or detection, but one of *creation*. And this is the source of the second deficiency of the valuation/detection story: it does not account for the need to search out options that do not exist yet. People cannot bargain over what does not exist, over that which they have not yet contemplated. And if innovation occurs primarily as a by-product of use and the result of accident, then the enforcement of use

restrictions may halt the innovation process *even before it has a chance to occur*. We have already seen actual examples of how user innovation occurs, and examined some reasons why users will have certain advantages in various types of innovation. However, why can't such advantages be duplicated by industry, via an organized, methodical program of research and development? This is where differential human capital comes in: it is the key constraint in the search for the higher-value uses of assets.

Human Capital as the Primary Constraint in the Search for New Uses (the Maximization Process)

Human capital, like other forms of capital, is a durable asset that produces income and productive outputs over long periods of time.[97] Unlike other forms of capital, however, human capital can never be separated from the individual who owns it (unlike physical capital - a manufacturing facility for example - which can be transferred to another individual). Human capital must also be distinguished from knowledge, which can be transferred from one individual to another. Human capital is thus a rival good, like a tangible. But it is also nontransferable, and so is more akin to tacit knowledge,[98] or to an inalienable asset. Over the years, human capital has assumed increasing importance in the literature on economic growth, with some commentators citing it as the single greatest source of economic wealth. With the upsurge of interest in the topic has come considerable divergence of opinion on what constitutes an appropriate definition of this important resource. Most would agree that the association of human capital only with certain types of investments – such as formal education and on-the-job training - represents an antiquated approach. Although skills and knowledge will always be important components of human capital, other factors, such as personal values and health also play a role in determining an individual's productive capacity. Thus, investments in health care, as well as the transmission of social norms that perpetuate values such as conscientiousness, or a 'hard work' ethic, also increase the stock of human capital. Along with an expanded appreciation of the types of competencies that constitute human capital, has come an expansion of the recognized sources of that capital. So, for example, we now know that school and work are just two of the institutions that impact our stock of human capital – the home and our community also do much to shape the attitudes, habits and beliefs that combine to mark the boundaries of our individual productive capacity. Indeed, some have proposed that human capital include "the entire life experience of the individual." The definition that will be used here builds on these expansive understandings, particularly those of the OECD[99] and NASA[100]:

> Human capital consists of the time, skills, knowledge, personality, capabilities, experiences, preferences, and perception of the individual. Human capital is acquired over the entire life cycle, and is used to produce goods, services or ideas in both market and non-market settings.

[97] See, e.g., Becker.
[98] Actually, tacit knowledge can be transferred, but that transfer is extremely difficult.
[99] Insert definition from Measuring What People Know
[100] Insert definition from SDI working group

This definition incorporates the salient features of previous definitions, but adds two new critical competencies: preferences and perception. These have been added both because their importance is evident from the case studies on user innovation, and because they have been identified as being of critical importance in the recent literature on innovation.[101] In general, preferences play a fairly uninteresting role in economic theory. They are taken as given, and are largely equated with consumer wants. They drive demand, to be sure, but beyond that their role is minor. Once we move away from the demand for, and consumption of, a non-durable good (e.g., I buy and eat an apple because I like apples more than bananas), however, something of a different order happens. Preferences begin to take on greater importance because they determine how individuals use and interact with the durable assets that they own. In other words, preferences determine what people actually *do* in every facet of their market and non-market lives, not just what they buy in the market.[102] Preferences in consumption might not appear to be that important in the search for new uses of assets, but preferences in use certainly are. And because these new uses create value via the combination of inputs (human capital, tangible assets and intangible assets) to create an output (the new use), they are difficult to distinguish from productive activity. It is thus reasonable to say that preferences do as much to guide production as consumption. Preferences, in other words, drive both our motivation to innovate, and the direction of that innovation.

Perception also has a pivotal impact on an individual's capacity to innovate, as it influences an individual's ability to perceive, and create, valuable opportunities in the environment. There are two types of perception, which are generally referred to as 'problem solving' and 'problem finding': 1) the first relates to the ability to perceive existing opportunities in the environment; 2) the second relates to the ability to see the world in new ways. The first largely solves known problems; the second finds both new problems and solutions. Perception thus affects our cognitive capacity to solve problems and develop innovative opportunities, whether they be old or new.

In organizations, differences in preference and perception across individuals can often be a barrier to efficiency, as they raise the cost of coordinating activity. So too, with governance rules in property ownership – if a purchaser of property has to get permission for a new use from the former owner, conflicting preferences and perception might block value-enhancing uses. But if we circumvent the coordination problem by replacing a governance rule with a rule of exclusion, then differences in preferences and perception can be seen in a new light. Instead of impediments, they now represent an unknown set of potential innovations. We know that the maximization process (the search for new uses) is always subject to constraints, but the cognition and motivation (perception and preferences) of the searcher are the two greatest constraints in the process. Thus, if we change the searcher, we *quite literally change the constraints*. What we therefore get when we move from governance to exclusion is the potential for a completely different search. This is why we can never ignore the issue of the scope of property rights (which is simply the 'initial allocation' problem restated), and why efficient breach has a place in both property and contract law. The relation between property rights scope (initial allocation) and innovative activity has been addressed by other commentators, of course. Some have argued that we will get innovations faster if we have narrower property rights, because different individuals will view the property differently. Others have argued that we will get innovations faster if we have broader property rights, because the incentive for speedy commercialization is increased

[101] See, e.g., Scott Page, The Difference.
[102] This observation is echoed in Nicholas Stern, Growth and Empowerment.

when we protect business models with servitudes and license restrictions. The argument here is related to the 'faster innovation' point, but is somewhat different. The assertion here is not just that we will get the same innovations faster, depending on who owns the property right; it is that we will get a completely different set of innovations, depending on who owns the property right. Indeed, we might miss out on some important innovations entirely. [The difference between this treatment and the Merges/Nelson treatment will be added here.]

[The difference between human capital and 'sticky information' will be added here.]

Now that we have seen how rules of exclusion foster user innovation, resulting in the efficient breach of governance rules, we can move on to discuss two main objections to the efficient breach theory. The first is that the *numerus clausus* has to be concerned with third party costs rather than first party costs, because the price mechanism accounts for the costs to parties within the chain of asset transfer. The second is that effective notice of use restrictions will resolve information asymmetries between the owner of the use rights and the possessor of the asset, and give rise to deals which result in the same pattern of innovative activity that would obtain under an efficient breach framework. The third party objection will be addressed in Part 3, and the notice objection will be addressed in Part 4.

Part 3: The Third Party Issue - From the Price Mechanism to Proxies

"If we knew how property would be used, the case for it would largely disappear. …Our faith in property *does not rest on the foreseeable results in particular circumstances* but on the belief that it will, on balance, release more forces for the good than for the bad."[103]

[103] Emphasis added. This is an adaptation of Hayek, The Constitution of Liberty: "If we knew how liberty would be used… . Our faith in liberty does not rest… ."

The Third Party Issue and the Price Mechanism

The first objection to the efficient breach theory of the *numerus clausus* arises from the fact that this theory focuses on the costs that novel property forms impose on individuals within the chain of transfer, or 'zone of privity'. Both the 'third party information costs' and the 'costs of verifying divided rights' theories of the *numerus clausus* are careful to focus on the information costs that novel forms of property impose on third parties.[104] The reason for this is straightforward: the costs to the persons within the 'zone of privity' (those to whom the property in question is transferred, either now or in the future) is generally considered to be accounted for by the price mechanism. That is to say, the more restrictions there are on a piece of property, the lower its price will be. Thus, the price will reflect all the costs that the idiosyncratic property form imposes on all those within the zone of privity. Any rationale for the *numerus clausus* must therefore (according to the canonical view) look outside the zone of privity for its justification, to the impact on third parties. But this view misses the very substance of what the *numerus clausus* is all about, what it is actually designed to *do*. What do I mean by this? Recall Coase's assertion that: "It can, I think, be assumed that the distinguishing mark of the firm is the supersession of the price mechanism."[105] Even if we consider the firm to be a 'nexus of contracts,' this only reflects the firm's internal complexity: ownership of the firm's assets remains hallmark of the firm (as this ownership literally provides the 'nexus'). It seems similarly straightforward to assert that property (a rule of exclusion) is nothing other than the supersession of the price mechanism, with the controlling mind of the firm owner replaced by the controlling mind of the property owner. If I own property, I don't have to pay others for the right to use it, and I don't have to ask others for their permission to use it. And when we consider user innovation, we see that the similarity is even greater: a property owner literally can be an entrepreneur. The *numerus clausus*, by replacing a governance rule with a rule of exclusion, thus operates to supersede the price mechanism: that is its purpose. If we accept the assertion that the hallmark of property is the supersession of the price mechanism, then the relevant question is not: Why does the *numerus clausus* exist despite the ability of the price mechanism to account for the costs to those within the zone of privity? Instead, the relevant question is: Why does the law favor (even force) the supersession of the price mechanism via the application of the *numerus clausus* principle? Like the theory of the firm, the theory of the *numerus clausus* forces us to ask: why supersede the price mechanism? Like in the theory of the firm, the answer is that there are costs to using the price mechanism, though in the case of user innovation, the costs center around the prevention of innovation rather than the determination of prices *per se* (though of course, when we consider disputes about paying for the permission to innovate, this is in part a pricing problem). And like in the theory of the firm, one of the major limits to the size of the firm is that: "the entrepreneur fails to place the factors of production in the uses where their value is greatest" and so it becomes more advantageous for another entrepreneur to take over the organizing function. Certainly, there are differences between Coase's formulation of the firm and property ownership. In this case, of course, the 'other entrepreneur' is not another firm, but the user-innovator. And Coase generally contemplated that the cost of organizing transactions within a firm would increase as the number of transactions increased, whereas here the limiting factor is not just the sheer number of transactions within the firm, but the cognitive capacity to engage in the innovating activity at all. Finally, the factors of production in user innovation are not the labor, capital, land and entrepreneurship found in the formal models; but rather labor, entrepreneurship and commodities. Technically speaking,

[104] Most (though not all) of the economic theories of servitudes share this characteristic.
[105] Nature of the Firm, in The Firm the Market and the Law at 36.

however, anything used in the production process is a factor of production, even if the formal models have to assume some elements away in order to be tractable.

Because the pricing issue engenders so much confusion, it demands that we examine it a bit more closely. It is undoubtedly true that property transfers occur within a framework of prices, just as the activities of a firm occur within a backdrop of pricing activity. The firm does not have to enter the product market, but it does have to enter the factor market, so what exactly does Coase mean when he refers to the price mechanism being superseded? Here we see why Coase's insight was so long in coming to the theory of the firm: for both the firm and for property, pricing in its transfer aspect often obscures the defining *lack* of pricing in the post-transfer aspect. Going forward, I will refer to pricing in its post-transfer aspect as 'use pricing', and to pricing in its transfer aspect as 'transfer pricing'. If 'property' is transferred in a transaction, this transfer is accomplished via 'transfer pricing' (an amount is paid to transfer ownership), but there is no 'use pricing' following that transfer (because ownership has been transferred, there is no ongoing need to compensate the previous owner). In other words, when Coase says that the price mechanism is superseded, he is referring to use pricing (the firm owns the residual of the factor of production), but not to transfer pricing (the firm had to pay to obtain that factor of production). In other words, the supersession of use pricing literally defines what it means to be an 'owner', or to be a firm, or to have 'property', or to have been party to a 'sale'.

Of course, confusion about the relation between pricing and property inevitably arises because the expected uses at the time of the sale[106] often factor into the price at which the property is 'sold'. If a seller suspects that a buyer will get greater use of the property, then he is more likely to ask for a higher price. Similarly, the price that the buyer is willing to pay can be expected increase as the value of his expected uses increase. But the question is *not* whether the price at the time of transfer (t1) accounts for the full expected value of the buyer (though it is reasonable to assume that there must be *some* difference between this and the actual price for a transfer to occur). The question is whether an *unexpected* change of value[107] some time *after* the transfer (at t2) will accrue to the buyer or the seller. Here we see the difference between transfer pricing (price at t1), and use pricing (price at t2). When property has been transferred, the price at t2 is always zero – hence the supersession of the price mechanism. If, on the other hand, no property is transferred in a transaction, then use pricing is not superseded, the price at t2 is greater than zero, and the endless permissions and payments inherent in a complete contracting (governance) scenario can be realized. Doubtless many readers will recognize that this view of property essentially equates it to ownership of the residual – that portion left over after all contract liabilities are settled.[108] But this definition, based on use pricing, also corresponds to the legal conception of property. The supersession of price, in either its transfer or its use aspect, generates a variety of rule types and organizational forms. These can be seen in Table 1.[109]

---

[106] Though more so in a bilateral monopoly situation than in a competitive market situation. This is why markets are so valuable: they widen the gap between transfer and use pricing, and are thus one of the best sources of rules of exclusion. See Part 5.

[107] This change in value could easily be a decrease, but for ease of discussion we will assume an increase in value (as the seller would obviously have no interest in trying to 'capture' (i.e., pay for) unexpected decreases in value post-transfer.

[108] In this respect, the view of Barzel et al is in full accord with the legal view of property. Where it differs is in its indifference to the transfer of the residual, which economists are generally willing to leave to the will of the initial owner.

[109] Actually, there is a fourth form as well. When we supersede both use and transfer pricing, we have one of three scenarios: uncompensated expropriation, theft or gift. But since the first two are of no significance in commercial transactions, we will not discuss them. The last, gift, is for all intents and purposes equivalent to a sale with a price of zero (in that the original owner consented to the permanent transfer but did not demand payment), so all discussion regarding Column D also applies to gift scenarios. See, e.g., UMG v. Augusto.

Table 1: Pricing Scenarios and the Resulting Organizational Forms/Rule types

	B	C	D
Resulting Form	License to Enter Real Property/Rental of Chattel	Continued Payments+Permissions after Permanent Transfer of Tangible/Servitudes (Governance)	Sale/Property/Firm (Exclusion)
Transfer Pricing	None	Transfer Pricing	Transfer pricing
Use Pricing	Use Pricing	Use Pricing	None

When we supersede neither use nor transfer pricing, we have a servitude or a governance rule[110] (Column C). Column C has been shaded because it represents highly contested ground in the law (but not in economics). At first glance, a servitude, which is property that has been transferred subject to one or more use restrictions, seems to be quite different from a governance rule, in which at least some, if not all, uses are subject to a scheme of payments and permissions from the transferor. But a use restriction is merely a use with an infinite price, so an absolute restriction on use is actually a paradigmatic example of use pricing. When we supersede transfer but not use pricing, we have what most people would recognize as a form of 'rental' (Column B).

When we supersede use pricing (the residual is transferred), but not transfer pricing (the transfer was compensated), we have a firm, property or sale scenario (Column D). Here we can also see a more precise definition of rules of exclusion. Exclusion is simply the supersession of use pricing: new uses of property *do not* have to be paid for under rules of exclusion; new uses of property *do* have to be paid for under governance rules. Only in Column D do we 'supersede the price mechanism', and so the terms 'supersede use pricing' and 'supersede the price mechanism' will be used interchangeably. Viewing property as the supersession of the price mechanism should be somewhat familiar to lawyers, who have long seen this issue in the context of sales law. In the words of Karl Llewellyn: "The law of Sales, as is well known, is in one phase part of the law of contract, in another phase part of the law of property."[111] In other words, a transfer of property is indeed priced (this is the contract portion of the sale), but if every use of a piece of property is priced and/or controlled by the 'seller', then no property has been transferred (the transfer of an unpriced portion is the defining characteristic of a sale). The response from the economic community to these continuing permission/payment scenarios (Column C) is generally that property rights are more fully specified. But of course, this means that the property rights of the *initial* property owner are more fully specified, while the property rights of the downstream 'user' are attenuated. The focus in economics is thus on the first owner in the chain of transfer, and any inability of this owner to create servitude-like restrictions, or demand repeated payments for new uses, is considered to be an attenuation of his rights. In other words, economics is generally indifferent to the choice between governance (Column C) and exclusion (Column D), as long as the transaction costs of creating and enforcing the governance rules are low.

The attitude of the law is radically different. The law of property is greatly concerned with mediating the 'property' rights of successive 'owners' in the chain of transfer, as the *numerus*

[110] Some have referred to the proliferation of these forms as a 'permission society'.
[111] Through Title to Contract and a Bit Beyond

clausus in fact creates such owners by legal fiat. Whenever the law interferes with the creation of a servitude or other governance rule (Column C), new uses of tangibles become unpriced (there is no use pricing), and the legal form is moved from Column C to Column D. Moving from Column C to Column D can be conceptualized in a variety of ways: as the supersession of the price mechanism, as a movement from governance to exclusion, as a restoration of full legal privileges, as a limitation on the number of property forms, or as a limitation on the ability to create divided rights.[112] However it is conceived, the practical result is that the law operates to force a series of partial property transfers from upstream owners to (what are now) downstream owners. The law's attitude toward the forms represented in Table 1 actually presents us with two puzzles. First, why does the law, as reflected in the *numerus clausus*, display such an antipathy for governance rules (Column C), and a concomitant preference for rules of exclusion (Column D)? Second, what exactly is the difference between Column B and Column D? In other words, why is tangibility so important in the law? We have already discussed the first puzzle in Parts 2 and 3: the law prefers rules of exclusion because they allow users to breach governance rules, thereby promoting user innovation. We will now discuss the second puzzle, the tangibility puzzle, as it is critical to the way in which the law implements its preference for rules of exclusion.

Why is Tangibility so Important in the Law?

I: Using Tangibility to Supersede the Price Mechanism

Let us return for a moment to the legal forms listed in Table 1. Column B includes arrangements in which a property owner does not permanently transfer property, but instead allows another person to use it temporarily. In the case of an immovable such as land, this involves temporary entry onto the premises (such as when I rent an apartment or enter a place of business); in the case of movable personal property, it involves temporary transfer of the item itself (such as when I rent a movie or carpet steamer and then return it). There is no doubt that the law of property would allow the owner virtually free reign in creating restrictions on, and requiring payment for, uses of an asset under Column B. But what exactly changes between B and D? In other words, what is the economic difference between license (or rental) and sale? If we put aside the legal consequences of the 'sale' label, we can see that the only real difference between license and sale is the physical fact of transfer itself. But the question arises: why does the initial owner have to give up what used to be hers, just because the physical location of the asset might have changed? Why does the ability to impose restrictions, and require ongoing payments, turn on the inconvenience of requiring that the property be physically returned? If we took the legal view that 'property is rights not things' to heart, then nothing of consequence should hinge on the act of physical transfer.

Tangibility makes Column B look quite different from Column D (in the one case, I possess it, in the other case, you do). This makes it relatively easy for possession to serve as a means of legal differentiation between the two. But as the tangible dimension is eroded, or as technology or notice enables control[113] after physical transfer, then the difference between B and D diminishes. Two things occur when this happens: 1) it becomes difficult to distinguish between license (B) and sale (D); and 2) intermediate forms (governance) emerge. Some examples from the intellectual property context may help to illustrate the point. Let's start with the simplest case. Suppose your neighbor leaves the local video store with a DVD. If he returns the DVD a week later, you know that he has

[112] Recall that I said the alternate definitions of the numerus clausus were complementary, rather than antagonistic.
[113] In the case of technology, control is *de facto*, in the case of notice, it is *de jure* (if enforced by the courts).

rented[114] the DVD. If he never returns the DVD, you know that he has purchased the DVD. You know this because of how he dealt with the physical copy. Now consider the application of the 'first sale' doctrine to the sale of a book. You enter a bookstore, pick up the latest Haruki Murakami offering (even though it's not in paperback yet) and bring it home. After reading it, you notice that there is a small legend printed on the inside: 'Licensed for personal use only. You may not resell, transfer or lend this book.' If you now want to sell your expensive hardcover and buy the latest Elizabeth George paperback, can you do this? Yes. Despite the attempt to characterize the transaction as a license by placing a line of print inside the cover,[115] this transaction is clearly a sale: you paid a single sum of money, and took permanent possession of the tangible object (the copy of the book). The first sale provision of the copyright act steps in and allows you to dispose of your copy of the copyrighted work.

Now suppose that instead of buying a physical copy of a book, you buy an electronic version for your fancy new e-book device. Suppose that every time you open the file, you are presented with a legend stating that this file is 'licensed, not sold', and that you may not resell it. You must click on 'I agree' to access your favorite book. Contract seems to be trying its best to interfere with your right to sell your copy. But your permanent physical possession of the file seems to support application of the first sale doctrine, despite the need to click a box to gain access to the file. In some ways, this looks like sale – you have permanent possession of the book file, and you don't have to make continuing payments for its use or possession. But the electronic nature of the copy (even though there are bits of data stored at a physical location on the device's hard drive) makes it harder to determine whether this is a license or a sale, because the technology has the capability to make your use of the book look more 'contract-like'. Now suppose that you purchase a new electronic book file, and this one automatically locks itself after you open it 10 times, unless you pay for an additional 10 viewings. Is this a license or a sale? Should it matter if you had notice of the restrictions, or when you were given notice of the restrictions (before you paid for the file or only after opening the file)? What if the file doesn't lock itself after 10 viewings, but instead *deletes* itself? These last 2 scenarios present us with 'pay per use' transactions, and they move us toward what some intellectual property scholars have called a 'permission society'. We will not discuss the appropriate legal categorization of these examples at this point, though we will return to some similar examples in Part 5. For now, the lesson to be taken from these examples is that tangibility does an incredible amount of work in the law: it not only provides a rough and ready way to justify the preference for exclusionary rules (i.e., the law can use the fact of physical transfer to force an asset from Column C into Column D); it also helps us differentiate between license and sale.[116]

Instead of allowing use restrictions (permissions) and payments to travel with a tangible asset, the law prefers to transfer standardized bundles of full legal privileges (full user). In other words, instead of allowing the first owner to maintain income and control over an asset after physical transfer to a downstream 'user', the law prefers to vest income and control in the 'user', thus creating a new owner. The former owner now has less than he would have had under a governance rule, while the new 'owner' now has more. Tangibility literally provides the law with a seemingly principled way to supersede the price mechanism (remember, new uses are unpriced under a rule of exclusion) between successive individuals in a transactional chain. I use the term 'seemingly' here because, as

[114] Which, of course, is a kind of licensing. We are assuming here that your neighbor is satisfied with the DVD (i.e., he is not returning a sale item), and that he is honest (i.e., he did not steal the rented DVD). Of course, even he tried to steal a rental, most stores now automatically convert rentals to sales if they are not returned within a certain period of time.

[115] Note that for the sake of simplicity, we ignore the complication that would be introduced if the notice was printed on the outside of the book and you had notice of it at the time of purchase.

[116] Tangibility (i.e., permanent transfer of the asset) is not the only factor to distinguish between license and sale, of course, but it is a necessary factor. [insert list of factors]

noted above, when we step back and question why the law draws such a sharp distinction between Column B and Column D, it is difficult to come up an answer that goes beyond the circular assertion that there has been a 'sale'. Tangibility is used to draw the distinction between B and D precisely because it creates an easily-locatable boundary between persons, and this boundary can be used to justify the supersession of the price mechanism. This is why notice is generally not as legally effective as a physical boundary in creating a new property form: notice operates in the exact opposite manner to tangibility. Notice operates to *prevent* the supersession of the price mechanism, as it allows the upstream owner to maintain post-transfer control over the asset. This control after physical transfer creates a situation in which all downstream uses can be priced, either per-use basis (in the case of a royalty or licensing fee), or infinitely priced (in the case of a use restriction). It should also be noted that the legal importance of tangibility, in addition to justifying the supersession of the price mechanism, has another effect: it allows individuals a significant amount of *freedom* in dividing property along its tangible dimension. After all, if I made a widget, and wanted to sell only half of it, the law would not step in and force me to sell the whole thing. I certainly could create this new property form called 'half a widget'. But a restriction on resale would not be looked upon so kindly: 'a whole widget that cannot be resold' would be sure to invite legal interference. Thus, the law allows significant customizability along the physical plane (the tangible), but not along the plane of 'use' (the intangible).

II: The Dimensions of Property

In many ways, economics shares the law's view that tangibles and intangibles should be treated differently. The tangible/intangible divide is roughly equivalent to the distinction between rival and nonrival goods. The canonical wisdom is that for rival goods, strong property rights are efficient. For nonrival goods, the situation is more complex, and involves balancing the threat of underproduction (when rights are too weak) with the threat of underuse (when rights are too strong). But what is a 'strong' property right? If we consider the question from the perspective of the initial owner, then it seems reasonable to equate strong property rights with fully specified rights in which he is free to craft any manner of governance rules. But this is not as straightforward as it seems. Consider a 'servitude' in which A transfers physical property into the hands of B, with a restriction that it only be used for a garden. We can view this situation in one of three ways: 1) we can consider it to be a case of divided rights in a single asset (B holds the right to use the land as a garden, while A holds the right to all other uses); 2) we can view it as a case in which a new property form, 'a garden servitude', has been created; or 3) we can view it as a situation in which B has a property right of narrow scope, and A has a property right of broad scope. Does this situation create 'strong' property rights? It all depends on whether you are A or B. Consider the feudal tenure system – that is considered to be the antithesis of private property rights. But if we replace the Crown with a private owner, and keep the governance rule in place (replacing services with money payments), we can see that the problem remains the same. There is something inherent in the legal notion of property that requires it to (at some point) transfer free and clear to another human being: property, at least in its legal incarnation, is an inherently dynamic construct. The reason for this is simple, though perhaps not obvious: there is no 'natural' or inherent scope of a property right, even in a tangible, and so the rule that rival goods require 'strong' property rights provides no real guide as to when use restrictions should, and should not be, legally enforced. Tangibility *appears* to give us a useful guide to property right scope, until we are faced with situations in which the dimension of use is separated from the dimension of physical space (possession). This can occur whenever rights in an asset are divided between multiple owners (as when a servitude gives one person physical possession and another

person the right to authorize specified uses), but it also occurs in situations of 'spillovers', which is simply the mirror image of the servitude problem. A common example of a spillover in the case of a tangible[117] occurs in situations of 'overlooking' – when we get use out of property that we do not occupy, by virtue of our proximity to it. Think about watching a drive-in movie from your rooftop, for example.

To get past the prescriptive impasse posed by property rights scope, a novel approach is required. What if we were to approach the question a bit differently, as if all property had both rival and nonrival dimensions? The question of how (and whether) intellectual property fits into the traditional notion of property has caused great debate in the legal community. But what if the right question is not how intellectual property is like property, but how property is like intellectual property? To see how this could be so, we need to break down property into its basic components. All property consists of three dimensions: a temporal dimension, a tangible dimension, and an intangible dimension. These can be seen in Table 2, below.

Table 2: The Dimensions of Property I

Intangible (use)	'Nonrival'
Tangible (space)	Rival (Thing, Embodiment, Copy)
Temporal (time)	Limited or Unlimited

We will not discuss the temporal dimension, although it is often quite significant in the realm of property. Instead, we will focus on the tangible and intangible dimensions. The tangible dimension in a physical 'thing' (a piece of land or a chattel) is obvious. But there is a tangible dimension in intellectual property as well. Much of the law of copyright and patent, for example, is concerned with injecting a tangible dimension into those arenas. The requirement of fixation in a 'tangible medium of expression', and the first sale doctrine, are two such examples from the law of copyright. The first sale doctrine separates the ownership of any particular 'copy' of the intellectual property (a book, e.g.) from the ownership of the copyright itself (a literary work, e.g.). It allows purchasers of copies to resell and use copies they have purchased, even though the copyright owner has the right to distribute his work in copies. The situation might appear to be somewhat different in patent law, since patent, unlike copyright, explicitly gives the patent owner the right, not just to make and sell, but also to *use* the patented invention. But here too, the law steps in and uses the tangible dimension to cut off the rights of the patent owner. Consider the most recent Supreme Court pronouncement on the issue of exhaustion, from *Quanta v. LG*: "The authorized sale of an article that substantially embodies a patent exhausts the patent holder's rights and prevents the patent holder from invoking patent law to control postsale use of the article." This ruling prevented the patent owner from

[117] Spillovers obviously become much more common as the tangible dimension of property is eroded.

collecting a licensing payment every time an article embodying the patent transferred hands. If we consider that each transfer could possibly represent a significant increase in value (depending on what use is made of the article[118]), we can see why the patent owner would want to capture that downstream value. The *Quanta* case represents a straightforward affirmation of 150 years of patent jurisprudence; however, it is particularly significant because it dealt with the application of the doctrine of exhaustion to *method* patents. Method patents are potentially much broader than other patents because they are not inherently linked to a device or thing. They literally claim a method of achieving a result - a process. It is easy to see why these sorts of patents would be problematic for the concept of exhaustion, which is tied to the sale of a tangible. Nonetheless, the Court once again chose to supersede use pricing between the patent owner and downstream users, by resorting to the device of tangibility. Once a method claim is embodied in any particular device, then the exhaustion doctrine takes effect, and the patent owner cannot demand payments or require permission for further use or transfer.

III: The Dimensions of Property and Human Capital

Why does the law, in the seemingly unrelated realms of real, personal and intellectual property, concern itself with locating tangible boundaries at which it can transform downstream 'users' into downstream 'owners' by superseding the price mechanism? To answer this, we need to take a closer look at the *intangible* dimension of property. We are used to thinking about intellectual property as a product of the mind, but all property is, in some sense, a product of the mind. When an owner seeks out new uses for tangible assets, for example, he applies his unique stock of human capital to the asset – his particular preferences, knowledge, abilities and skills – generating something which has both tangible and informational dimensions. Sometimes the informational component appears so significant, and the human capital component so specialized, that we carve it out and call it 'intellectual property'. This tends to obscure the tangible dimension of intellectual property. At other times, the information generated by the new use, and the human capital required to generate that new use, appears more mundane, and we are content to see it as nothing more than a new 'use' of a 'thing'. This tends to obscure the intangible dimension of physical property. Nonetheless, the two cases are more alike than they first seem: the search for new uses of tangibles requires the application of human capital and produces information, just as the creation of 'intellectual property' does. The intimate link between human capital, information and property can perhaps better be seen if we consider all the possible combinations and their resulting legal classifications. See Table 2, below. The most important classifications, at least for our purposes, have already been discussed: things, information and IP. Inalienables are included for completeness only, but as a category they have been studied extensively by Radin, and include things (such as body parts) which protect or express a personality interest of the individual.

[118] This case involved an article that was used as a component in downstream products, so the increase in value could be particularly large. However, the principle does not depend on the distinction between a 'component' and an 'end product'. In the case of end products, increases in value can also be quite large, whether because new uses are made of them, or whether the market demand for them rises.

Table 2: Information, Human Capital and Legal Classification

Information Content	Human Capital Content	Legal Category
Low	Low	'Thing'
Low	High	Inalienable
High	Low	Information
High	High	IP

Looking at property this way creates an intangible dimension that is more complex than the standard rival/nonrival distinction. Instead, we have an intangible (use) dimension that has both rival (human capital) and nonrival (information) components. Consider the shaded area in Table 3, which provides a more detailed breakdown of the intangible dimension of property. The dimension of use is complex, not only because it has both rival and nonrival components, but also because it consists of a complex interplay between these components. This interplay is complex because each serves as an input for, and an output from, the other: processing information increases our human capital, while the operation of human capital on the environment produces information.

Table 3: The Dimensions of Property II

Intangible (use)	Rival, Nontransferable (Human Capital)
	Nonrival, Transferable (Information)
Tangible (space)	Rival, Transferable (Thing, Embodiment, Copy)
Temporal (time)	Limited or Unlimited

So now we can see that the intangible dimension of property has both rival and nonrival components. But what does tell us about the role played by tangibility in the law? By using the *numerus clausus* to supersede the price mechanism across boundaries of tangibility (i.e., when the tangible permanently changes hands), the law allows individuals to interact freely with assets, without the need to pay for (or get permission for) any new uses that may result. As we have already seen in Part 2, what any one individual will make of this freedom will depend on his stock of human capital – his attitudes, preferences, knowledge, abilities, and the like. But we cannot determine *ex ante* whether any one individual will either have, or will use, the human capital which will produce new uses of resources. What we do know, however, is that each person's human capital is unique, so that when a tangible asset changes hands, there is a potential opportunity for new uses to be generated. By superseding the price mechanism across planes of tangibility, the law thus uses tangibility as a very rough proxy for human capital. There are three things we know about human capital that make this strategy a wise one: 1) human capital is a critical input to the creation of new uses; 2) human capital is rival and non-transferable; and 3) each person's stock of human capital will be unique. The use of tangibility as a rough proxy for human capital results in property being organized around its rival, nontransferable dimension. From an efficiency perspective, this makes sense: the downside of getting an initial allocation of rights 'wrong' is much greater when a rival, nontransferable input is at stake. It is much easier to cure incorrect allocations when inputs are nonrival (these are inherently leaky), or rival and transferable (private parties can transfer them). The use of proxies is ubiquitous in economic activity; they are employed whenever direct measurement of a relevant characteristic is difficult or impossible (i.e., costly). A simple example occurs when purchasing fresh fruit: we use easily observable characteristics, such as color and firmness, to substitute for measurement of what is actually relevant to us - the flavor of the fruit. The use of proxies has its own cost, however, as all proxies suffer from some degree of inaccuracy. In the case of fruit, for example, breeding often produces varieties that are highly colored yet low in flavor. In this case, a consumer would want a proxy that more accurately measures the flavor profile, and might start using the size of the fruit, instead of the color, as a proxy for flavor (smaller fruit varieties often have more intense flavor). Here is where the use of proxies in property law differs from the use of proxies elsewhere in economic life. Unlike in most cases, we do *not* want the 'tangibility as stand-in for human capital' proxy to become more accurate. If we did, the specter of governance rules replacing rules of exclusion would not present a problem (and there would certainly be no need for the *numerus clausus*).

The Enduring Logic of Rough Proxies

We know that each person's stock of human capital is unique, and that each person therefore brings a unique set of inputs to the process of constrained maximization. But there is a problem here: we cannot predict who will have the stock of human capital that will produce a new use (in other words, we do not know who will have the cognitive capacity). And even if a person does have the cognitive capacity, we cannot predict whether any particular individual will use it (that is to say, we do not know who will be motivated to search for a new use). This poses quite the problem for property law: how do we design a system that enables us to find what we don't even know that we are looking for, or who might find it? Furthermore, we also saw that accident plays a large role in generating innovative outcomes, so how do we deal with that aspect of reality? How, in other words, do we design a system of property rights around fundamental uncertainty and luck?

Economists have long been aware that imperfections in cognition and motivation pose a challenge to the concept of maximization.[119] How can we usefully speak about 'maximization' when the foresight (i.e., cognitive capacity) of individuals is imperfect, and their motivations are often wildly divergent and stubbornly non-maximizing? The canonical answer to this problem was worked out by Armen Alchian in 1950: "Realized positive profits, not maximum profits, are the mark of success and viability. *It does not matter through what process of reasoning or motivation such success was achieved*."[120] (Emphasis added.) In other words, even a random set of individuals, who are completely lacking in foresight and who are not motivated to maximize profits or find the 'best' use of resources, will be subject to the selection mechanism of the market. Over time, the environment itself (the market) chooses successes and weeds out failures. Those who realize some positive profit will survive, while those do not will fail. Success is its own reward, in other words, even if success was neither the aim nor the goal.

This treatment of uncertainty is reminiscent of the 'tale of the typing monkeys'[121]: if we put an infinite number of monkeys in front of typewriters, one of them will eventually produce a readable work.[122] In other words, randomness, when coupled with trial and error, will eventually produce order. But, of course, there is a key variable here: the number of monkeys. More monkeys means more trial and error, and the "society that permits the maximum generation of trials will be most likely to solve problems through time."[123] Here is where the law's treatment of uncertainty differs from the standard economic account: the law is sensitive to the fact that the scope of property rights will impact both the *number* and the *type* of trials that will occur. It may very well be the case that the environment chooses successes over failures, but the scope of property rights will help determine what options the environment has to choose from. More monkeys mean more options; different monkeys mean different options; better monkeys mean better options. In economics, the term 'uncertainty' is generally used to refer to the fact that we do not know what the possible states of the world will be.[124] But what those states actually turn out to be will depend on who conducts the search, and on how the results of those searches arise out of random events. In many ways, the term 'search' does not quite capture what happens when individuals seek out new uses of assets. They are not just searching the environment, they are acting upon it, changing it - literally creating it – and this process is highly context sensitive.

Uncertainty introduces an inevitable trade-off into the process of constrained maximization. On the one hand, uncertainty is a negative factor in human experience and business planning because it makes it impossible to identify or predict wealth maximizing opportunities, and to plan the course of action most likely to achieve our goals. Uncertainty that goes beyond mere lack of information implies that hit-and-miss experimentation is necessary, and experimentation is inevitably wasteful.[125] Some attempts will fail miserably, while others will result in unpredicted (and unpredictable) success. So the flip side of uncertainty's inevitable waste is its potential for unpredictable success. It therefore represents a potential upside – the value of which cannot be measured in advance. The potential upside of uncertainty is perhaps better captured by the term complexity instead of

---

[119] See, e.g., Tintner, Theory of Choice, Pure Theory of Production

[120] Alchian, 213.

[121] Which is to be distinguished from the 'British Museum Algorithm'. See, e.g., Campbell.

[122] See, e.g., Nassim Taleb.

[123] North at 81 (Speaking of Hayek.). See also, The Constitution of Liberty.

[124] The term uncertainty can refer to either 1) the situation in which we know what the future states of the world will be, but do not know their probabilities, or 2) the situation in which we do not even know what those states will be. The closely related term 'risk' refers to those situations in which we can assign probabilities to the possible future states of the world.

[125] Demsetz, Nelson.

'uncertainty'. The existence of persistent complexity in the environment implies that it cannot (and should not) be reduced, but instead must be explored. Perhaps an even better term than 'complexity' is 'luck.' Luck will always play a role in economic life: we all begin with different endowments of human capital (genetics is the purest form of luck), and we all encounter chance events which fundamentally alter our economic fortunes, whether for the good or bad. Despite the unfairness of it all, and despite the well-meaning but misguided musings of legal scholars,[126] we would not want to reduce the role that lucks plays in our lives. Not only would it be incredibly expensive to do so,[127] it would rob us of one of the best sources of economic wealth: "the reason free markets work is because they allow people to be lucky, thanks to aggressive trial and error, not by giving them rewards or 'incentives' for skill."[128]

Dealing effectively with uncertainty and luck is where property law's use of rough proxies makes an enduring contribution to the process of constrained maximization. It is true that innovation (the search for highest-value uses of assets) is an inherently unpredictable and context-sensitive process. We cannot identify ahead of time who will innovate, or what that innovation will be. But there *is* one thing that we know for certain when a tangible asset is permanently transferred from one individual to another (let's say from you to me). We know that *you are not me*. We know that your stock of human capital is different than mine, and therefore there is the potential for a whole new set of value-enhancing and information-generating new uses. By using the permanent transfer of a tangible as a very rough proxy for human capital, by incorporating the concept of efficient breach into the very fabric of property law, the *numerus clausus* principle facilitates the search for new uses of assets, thus optimizing the process of constrained maximization over time. The rough proxy of tangibility performs two related functions in property law: 1) it increases the number of trials; and 2) it randomizes who will perform them. The first function addresses the uncertainty problem by increasing the likelihood of success over time. The second function helps to generate luck, by increasing the chance that we will match the unexpectedly 'right' stock of human capital, with the unpredictably 'right' asset. Below is a series of schematic representations of how using tangibility as a rough proxy in property law affects asset use. In each scenario, we begin with the simplifying assumptions that each individual has only one possible use of an asset, and that each individual's use is unique.

In Figure 1 (next page), we consider the simple case of a unique chattel, i.e., of which there is only one copy (a piece of land would also fit the scenario). R_0 refers to the asset in the hands of some original owner – for the sake of discussion, we can ignore what the use of the 'resource' was before transfer. At T_1 the chattel is sold subject to a governance rule – for example, that it can only be used for a specific purpose. If the governance rule 'travels with the asset', then asset use is the same over time, no matter how many times the asset is transferred in the future. In figure 2, the same asset is sold subject to a rule of exclusion. Here we see that asset use, even with a single chattel, can then vary each time the asset is transferred.

[126] Haifa conference
[127] Cheung
[128] Taleb at xxi.

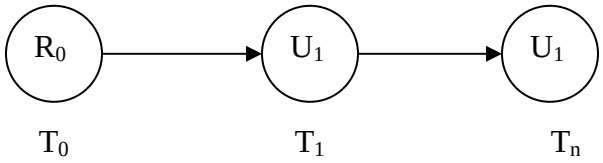

Figure 1: Asset Use Under a Governance Rule – Land, Unique Chattel

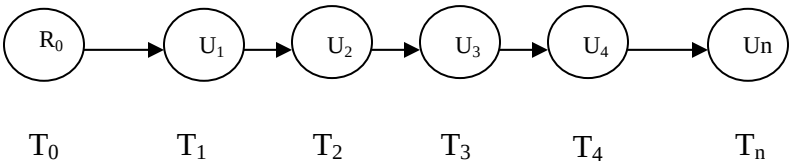

Figure 2: Asset Use Under a Rule of Exclusion – Land, Unique Chattel

In Figure 3 (next page), we move on to the more complex case of a non-unique chattel, i.e., of which there is more than one copy. Here, we also include the case of intellectual property – as the IP can also be embedded in multiple copies of the work. Once again, we begin with the case in which the chattel or copy is sold subject to a governance rule that states that the asset can only be used for a specific purpose, and the governance rule runs with the asset. As before, we can see that asset use is invariant over time. In Figure 4 (also next page), we move on to the case in which a non-unique chattel or IP embodiment is sold subject to a rule of exclusion. As before, asset use can now vary each time a chattel or embodiment is transferred. We assume for the sake of simplicity that each individual has a unique use that does not vary with time. If this is not the case (and it is likely not to be), then the potential variance in asset use under a rule of exclusion is even great than illustrated in Figure 4.

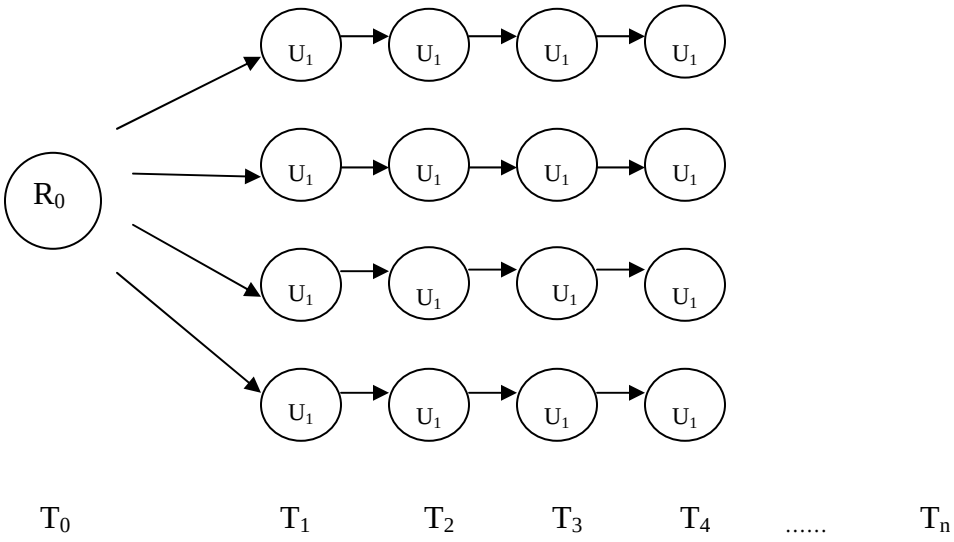

Figure 3: Asset Use Under a Governance Rule – IP, Non-Unique Chattel

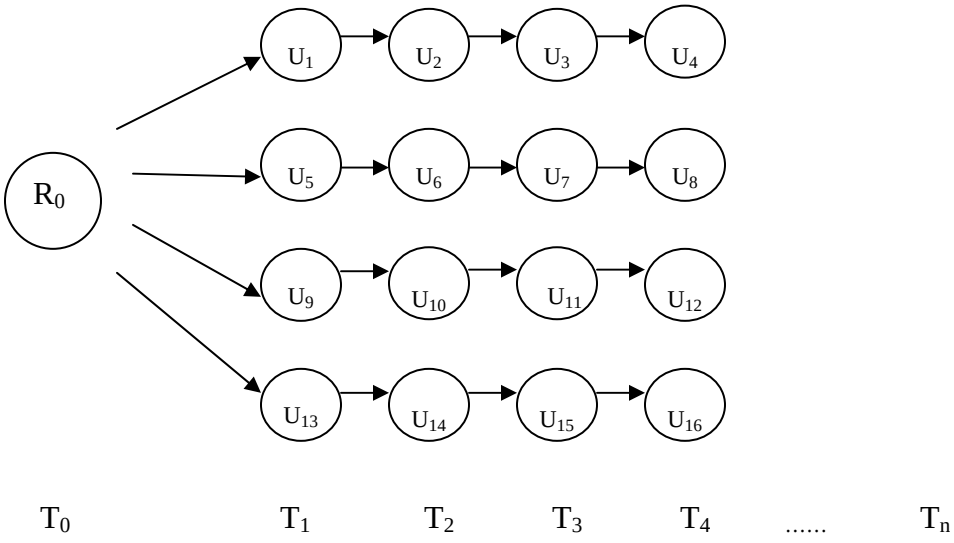

Figure 4: Asset Use Under a Rule of Exclusion – IP, Non-Unique Chattel

We have already touched upon the conceptual similarities between chattels and intellectual property, but these figures illustrate that there is an economic similarity as well. In other words, *there is an underuse problem with tangibles as well as with intangibles*. Intellectual property, because it is non-rival, is often touted for its 'multiplicative effect' – because it can be used by many people at once, it has a large downstream potential for generating downstream variations. Granting an IP right can block these variations, and thus poses the danger of underuse. But here we can see that chattels have a similar multiplicative effect. Allowing use restrictions on chattels thus poses a similar danger of underuse. Although the scale of the effect is greater in IP than with chattels, the underlying dynamic is the same. By assuming that every time a tangible asset is transferred unique human capital has the potential to generate a new use, the law promotes both intentional and accidental innovation. This is a powerful device for optimizing the process of constrained maximization. The assumption that each person *will* generate a new use is, of course, a simplification. An actual pattern of user innovation might look more like this:

[I am having trouble with these diagrams, this diagram will be inserted later.]

Here we see that even if many, if not most, users do not innovate, there will still be more innovation in a regime in which users operate under rules of exclusion with respect to their assets, than one in which they are subject to use restrictions (governance rules).

Now that we have taken a closer look at how the law replaces governance rules with rules of exclusion, and have answered the first objection to the efficient breach theory of the *numerus clausus,* it is time to move on to the notice controversy. To what extent does notice alleviate the concerns that animate the *numerus clausus*? The answer to this question will be explored in Part 4.

Part 4: Notice Does Not Cure Asymmetries in Cognition and Motivation

"Every assignee of a lease has notice of the lessor's covenants; consequently no covenant, however absurd, could be made by a lessee, that would not of necessity run with the land in equity into whose hands soever the land might come, and, all the decisions that have been made by the courts with respect to such covenants being collateral or in gross, would be of no avail."[129]

[129] Keppell v. Bailey.

I: The Notice Issue

The second major objection to the efficient breach theory of the *numerus clausus* is that, in this theory, the availability of new property forms does not vary with the cost of providing notice of those forms. This is in marked contrast to the 'third party information costs' and the 'costs of verifying divided rights' theories of the *numerus clausus*. Because effective notice significantly reduces (and often completely eliminates) the costs that third parties must incur to determine the characteristics of novel property forms, notice plays a prominent role in both of these theories. But notice also reduces the costs to parties *within* the chain of property transfer as well, as effective notice can act as a replacement for contractual agreement between asset transferees. For this reason, debates about the appropriate legal ramifications of notice have been waged in property law for a long time. They have been particularly pointed in the literature on servitudes, with some commentators asserting that both the principles of private property ownership, and freedom of contract, call for a regime in which effective notice is the *only* requirement for the creation of novel servitudes. Despite the forcefulness of these debates, the law continues to resist the idea that notice alone should draw the line between permissible governance rules and legally imposed rules of exclusion. And in fact, the judiciary has always been aware of the challenge that notice poses for the law's attempt to place limits on transactions involving property: "So a person who had conveyed land, and subjected it to covenants in the hands of his vendee, could at once make sure of those burdens following it into the hands of all holders to whom it might pass by taking the precaution of notifying the covenants *in some effectual though easy manner*, as by publication in some place near the premises where the purchaser must needs observe the announcement." (Emphasis added.) These words from Keppell v. Bailey are almost prescient in their anticipation of the conundrum that notice poses in the digital age: If we can cheaply and easily provide notice of unusual restrictions on real, personal and intellectual property, then do the legal limitations built into each of these regimes have any continued purpose?

Implicit in the assertion that effective notice should provide the only barrier to the creation of novel servitudes (i.e., governance rules) is a belief that the only problems created by such arrangements are problems of information asymmetry. If this is true, then once notice cures the information asymmetry, there is no reason for further concern. This view has recently been used by Antony Dnes and Dean Lueck to construct an information asymmetry theory of servitude law, for example: "The recurring worry of courts is that 'novel restrictions' could so encumber the land that the market for land would be undermined. Actually, *this cannot happen if there is clear information about the restrictions*, because they would be reflected in market value and sellers would soon learn the extent of their financial loss from high levels of restriction."[130] (Emphasis added.) Doubtless, this is correct: if information asymmetries are the only cause for concern in the *numerus clausus*, then notice should take care of things quite nicely. Notice will provide market participants with information about restrictions, this information will result in changes in the price of burdened assets, and the market will work to rid us of any unwanted restrictions. But the assertion here is that asymmetries in motivation and cognition, and not merely asymmetries in information, are the real concern. The differences between the standard information asymmetry account, and the cognition and motivation asymmetry account ('CM account'), are summarized in Table 6, below.

[130] Dnes, p 14.

Table 6: Types of Transactional Asymmetries and the Impact of Notice

	A's State	B's State	Information Asymmetry?	Transactional Hindrance	Cured by Notice?
1	Information	Information	No	Motivation to Deal	No
2	Information	Ø	Yes	Measurement	Yes*
3	Ø	Information	Yes	Measurement	Yes*
4	Ø	Ø	No	Cognition, Motivation to Search	No

* Even after information is produced, and notice cures an information asymmetry, a motivation to deal asymmetry may be present.

There are two principle differences between information asymmetries and the kind of asymmetries at issue in this theory. First, information asymmetries are cured by notice; whereas cognitive and motivation asymmetries are not. Second, each type of asymmetry implicates a different set of problems. Information asymmetries present us with problems of measuring what already exists; whereas cognitive asymmetries present problems of producing what does not yet exist. Motivation asymmetries present problems of both producing what does not yet exist, and dealing over what has already been measured. Looking at Table 6, we can see that Row 4 deals with what happens *before* information about new uses is produced, and Row 1 deals with what happens *after* that information is produced. From Row 4, we can see that in the case of cognitive asymmetries (and asymmetries in motivations that affect search[131]), neither party has present information about new uses. However, a unique set of perceptions and preferences, when combined with an ability to search (in the case of intentional innovation), or at least with an unfettered freedom to interact with the asset (in the case of accidental innovation), may produce them at some future time. From Row 1, we can see that in the case of asymmetries in the motivation to deal, both parties now have full information about the possible new uses of the asset, but the party who owns the right to authorize the new use refuses to give permission. The differences between the information asymmetry account and the CM account reveal two unexamined assumptions behind the information asymmetry view. First, there is an assumption that the information produced *will be the same* regardless of who owns the right to determine the new uses of the property (information production). Second, there is an assumption that once information about higher-value uses is revealed, parties will be motivated to strike bargains that allow those uses (motivation to deal). Neither of these assumptions is necessarily valid.

[131] Motivation asymmetries affect information production because they determine whether we are motivated to search, and they influence the direction of search. See Part 2. This is separate from the motivation to deal, which only comes into play after information has already been produced.

II: Information Production

We have already addressed, in Parts 2 and 3, the primary reason why the information produced from (and about) asset use will tend to vary depending on how rights are initially allocated: differential human capital. Each person's stock of human capital is unique, and this means that each person has a unique perception of the world around him, and a unique set of preferences about that world. Differences in individual perception create persistent cognitive asymmetries across individuals. Each individual's cognitive capacity is unique, and this affects their ability both to solve existing problems and find new ones. Each person's search for new uses of assets will therefore be unique, and the information that results from those searches will also be unique. Our preferences also have an impact on information production. They affect both our motivation to search, and the direction that search will take. In part, the motivation to search will be intrinsic – that is to say, it will arise from your internal set of preferences, rather than the external incentive of monetary[132] return (or disincentive of cost). This sort of motivation, which was also discussed in Parts 2 and 3, will always vary from person to person, and as such it is a persistent asymmetry, much like cognition.

In addition to intrinsic motivations to search, however, there are also extrinsic motivations to search. That is to say, your motivation to search will be reduced if you know you will have to pay for that search (at least if your preferences respond to price). Even if permission was automatic, and even if payments were costless to determine, administer and enforce, their very presence reduces the incentive of the downstream party to search for new uses of an asset. In other words, we limit property rights for the very same reason that we grant them in the first place – to provide the incentive for the possessor of the asset to invest in its development. The strength of these extrinsic motivations will vary depending on the results of the search: whether I find a new use that is merely personal, one that is potentially marketable, or find nothing at all. In those cases in which the new use is potentially marketable (rather than of merely personal benefit), there may be enough potential profit to search and then deal (in which case we are still subject to the refusal to deal problems discussed below). But it may be impossible to know ahead of time whether I'll just have to pay more for a personal use, or whether I will generate a use that I might be able to parlay into a joint profit opportunity (at which point I may still be subject to Arrow's information paradox, and the valuation problems canvassed by the standard valuation/detection account that was discussed in Part 2). There is a third possibility as well: I may search and simply find no new uses at all. Depending on the type of restrictions and type of asset at issue, this still might require me to pay for that failed search. The uncertainty inherent in these three scenarios (pay for personal use, possibly split the profits for a marketable use, or pay for a failed search), has not been directly examined by other commentators. However, it is somewhat similar to some existing property rights theory developed by Robert Merges (IP should be granted where searches are highly uncertain) and Henry Smith (property rules allow the owner to unilaterally pick out uses because the future value of resources is uncertain). For this reason, the extrinsic motivation to search will not be discussed any further.

[132] Actually, the return need not be money *per se*, it simply need be some external reward.

III: Motivation to Deal

Unlike cognitive asymmetries, which in this framework impact only information production,[133] motivation asymmetries impact both the production of information, and the subsequent bargaining over that information. A person's *motivation to search* for new uses of assets will help determine whether there are any new uses worth striking deals over. His *motivation to deal* over those new uses will determine whether these deals are in fact struck. In general, scholars who suggest that the *numerus clausus* should turn on the costs of providing notice have not recognized that the first (the information production) problem exists. Further, they generally assume that the second (the motivation to deal) problem is not really worth considering. The reason that scholars ignore the second problem is straightforward. Although notice generally operates to *unilaterally* bind downstream parties to restrictions, there is an underlying assumption that *if* it were economically rational for parties to come together and alter the restrictions, they would indeed do so. That is to say, the prospect of being better off will be sufficient to motivate individuals to strike deals. This can be seen quite clearly in the quote from Dnes and Lueck's servitude theory (price changes will motivate sellers to alter restrictions), but the assumption is also woven deep into the fiber of economics. Nevertheless, disputes about an individual's motivation to deal have arisen now and then even among those who tend the economic cannon. This is best illustrated by the sharp series of exchanges about parties' willingness to deal between Ronald Coase and Paul Samuelson, which prompted the following from Coase: "[those] who find it impossible to conclude agreements will find that they neither buy nor sell and consequently will have no income. Traits which lead to such an outcome have little survival value, and we may assume (certainly I do) that normally human beings do not possess them."[134] The debate about motivation to deal in situations of bilateral monopoly spilled over into the legal academy, and was the spur for a probing re-examination of the Coase theorem. That debate will not be revisited here. Instead, we will focus on the two 'motivation to deal' issues that are present in the search for new uses of assets: 1) the motivation of the searcher to disclose the new use; and 2) the motivation of the owner of the use rights to grant permission for the new use.

a) The Motivation to Disclose

[This section is in progress.]

This leads us to the fourth assertion about user innovation:

> User innovation will be more likely to be revealed (freely or otherwise) if users are given free and full use of the assets they possess.

b) The Motivation to Give Permission

Even if a user reveals the successful results of his intentional or accidental search, the owner of the use rights may refuse to conclude a deal that gives permission for that new use, even if that deal would make everyone better off. Although such behavior flies in the face of the

[133] That is to say, once information is produced and is made available via notice, it is assumed that there is no cognitive barrier to interpreting that information. Such a cognitive barrier does appear in Smith & Merrill's theory (too much information produces noise that makes interpretation difficult).
[134] Coase, Notes, 161-162.

rational actor model, it is far more common than often thought. We will focus on four situations in which the motivation of real economic actors has hindered the development of efficient transactions[135]: 1) preferences that do not respond to price; 2) refusals based on the identity of the requester, and not the content of the requested permission; 3) the tendency of entrenched economic actors to overestimate the value of old ways of doing things, and underestimate the viability of new ways of doing things (incumbent myopia); and 4) endowment effects.

i) Preferences That Do Not Respond to Price:

If sellers of assets always responded rationally to price, then there would be little reason for the *numerus clausus*. Why is this so? The reason is simple: changes in price would motivate sellers to impose only those restrictions (i.e., create only those new property forms) which were desirable to prospective buyers. If a buyer burdened a small lot of land with a covenant preventing the removal or repainting of a Day-Glo pink polka-dotted house, for example, this would make the land less valuable to most (if not all) prospective buyers. Idiosyncratic restrictions will put severe downward pressure on the price of an asset, and the seller will respond rationally by lifting the restrictions to make the asset more marketable. In this way, the market will weed out restrictions that are of little economic value, so it is, in theory, inappropriate and unnecessary for the law to interfere and second guess the market. It is easy to see why this view is so persuasive: not only does the market make a judgment call that seems inherently difficult for judges to make (how can judges determine which restrictions are valuable and which are not), it does so in a way that affirms individual choice and freedom. Certainly, the market does much to weed out undesirable products from the market (does anyone even remember that 'I Hate Peas' existed briefly in the 70's?). But asset restrictions don't seem to be as responsive to consumer preferences: asset restrictions often become standardized across industries and across asset types, and this makes them much less responsive to price pressures.[136] These situations, in which prices do not respond to buyer preferences, will be discussed in Part 5. For now, we will consider the other side of the issue: what happens if sellers do not respond to price changes? That is to say, what if their preferences do not respond to price? When this happens, a seller may be willing to accept a very low price for a burdened asset (as long as the burden is maintained). Equivalently, a seller may be unwilling to give permission for the burden to be lifted from the asset – for any amount of money. Indeed, it is accurate to say that, in some cases, the denial of permission actually prevents a new asset from coming into existence (or results in its destruction). We will discuss two canonical examples of price-resistant preferences, one from intellectual property and one from real property: parody and conservation easements. One common example in which preferences pose a pricing problem will *not* be discussed: expropriation of housing and land. These situations are taken to be cases of undercompensation, rather than cases for which no amount of compensation would be sufficient.[137]

Parody: [This section is in progress.]

Conservation easements: [This section is in progress.]

---

[135] There may, of course, be more categories. But these are taken to be those that are of particular importance.

[136] Problems of standardization in asset restrictions, whether it be in the movie industry, the insurance industry or in software licensing, e.g., will be discussed in Part 5. At that time, we will also explore the issue of thin versus thick markets (i.e., do *numerus clausus* concerns apply only in the case of thin markets and/or monopoly. Not surprisingly, it will be asserted that *numerus clausus* concerns apply in all markets.

[137] There is a rich literature on the chronic undercompensation problem in expropriation cases, which discusses how subjective valuations of homes and family land can result in an uncompensated increment, and how the appropriate division of the surplus in eminent domain cases is particularly difficult to determine. [Cites will be added.]

Price-resistant preferences are extremely problematic for economic theory. On the one hand, it seems reasonable for economics to refuse to engage in interpersonal comparisons of utility, and to ignore the wealth effect. There doesn't seem to be any other way to make economic analysis tractable. On the other hand, there has to be an assumption that price has *some* correlation to the relative value of actual uses. Otherwise, relative efficiency has no meaning. But if individual preferences do not respond to price, there is no way to rank uses, and no way to judge whether assets are moving to their 'higher-value' uses. In other words, if I can't be bought out at any price, then my price is infinite.[138] Does this mean that my use is the higher-value use? Probably not. And even if my price-resistant use does happen to be the 'higher-value' use (though how would we know for sure), it would certainly not be because my buy-out price happened to be infinite. If we do not know when resources are moving to their higher-value uses, then how do we know when efficiency has been improved? Economists are not completely unaware of this problem (even though formal models generally assume it away): "allowance for noneconomic motivation suggests that the 'natural' pressures toward efficiency represented by agreements, mergers or contractual arrangements generally among affected parties may be much less effective than the formal analysis seems to imply. The genuine zealot … may be insensitive to proffered compensations."[139] This recognition of the problem of price-resistant preferences is heartening, but it may underestimate its prevalence. Evocation of the 'zealot' may be comforting, at least for purveyors of economic theory, because it implies that the problem is a relatively rare and unusual one. After all, who of us, in the ordinary course of things, can afford the price of unadulterated zeal? But what if zealotry were not a rare wrinkle in economic theory, but rather a common human characteristic? Perhaps we have a bit of the zealot in all of us, and perhaps it simply takes the right set of circumstances (i.e., type of asset) to bring it out. Indeed, perhaps it sometimes takes the right *person* to bring it out. Which brings us to our second category of price-resistant preferences.

ii) Refusals Based on the Identity of the Requester, Not the Content of the Requested Permission:

The first category dealt with situations in which the *content* of the requested permission generated refusals that were invariant with price. This category deals with those cases in which refusals to license depend, not on the content of the permission, but on the *identity* of the individual requesting that permission. If the owner of the use right is responsive to price, then the identity of the person requesting the permission should not itself matter, so long as the individual possesses whatever skill and competence is required to undertake the requested action. If permission is withheld because of personal feelings towards the requesting party, on the other hand, then something very different is going on. One much-ruminated example of such a refusal to license comes to us from the aviation industry: the Wright brothers' refusal[140] to license the best and most profitable aircraft manufacturer of their day. Now, a caveat of sorts is in order, because Orville and Wilbur were not the easiest men to deal with, for *any* prospective licensee. The Wrights were slow to license as a general matter, and there were several reasons for this. First, they were extremely worried about disclosing their technology during negotiations, oftentimes refusing to demonstrate a

[138] Not being able to buy at any price is very different than not being able to sell at any price. Not being able to buy at any price results in an infinite price. Not being able to sell at any price results in a price of zero. Being willing to sell at a price of zero, however

[139] Buchanan, AC at 12.

[140] Actually, this seems to be more the work of one brother, Orville Wright. While he was alive, Wilbur seemed more than willing to license Curtiss. [Insert quote from telegram.] Though Wilbur also drove a hard bargain, due to overconfidence, see fn 14.

flying aircraft until after a contract was signed (in many cases, this meant that no contract was signed at all).[141] In part this reflected a concern about revealing their innovations to potential imitators – a classic Arrow's information paradox problem. This concern seems somewhat overstated, however: it is not clear that a potential competitor would learn anything that was not already disclosed in public documents from viewing a flight demonstration from a reasonable distance. In any event, in some cases, the Wrights seemed to be offended that a prospective buyer would even question that they had built a viable flying aircraft. Second, the Wright's believed that they had such a head start on other innovators that it would take at least five years for a competitor to independently solve the flying problem: "If it were indeed true that others would be flying within a year or two [from 1906], there would be reason in selling at any price but we are convinced that no one will be able to develop a practical flyer within five years."[142] This overconfidence exacerbated the unreasonableness of their demands during negotiations: the royalties they demanded were too high, and they were willing to leave the table at the drop of a hat.

Despite this general reticence, the Wright brothers did eventually license various entities[143], but there was one individual whom Orville steadfastly refused to license: Glenn Curtiss.[144] "Wright… announced that he will consider lenient royalty arrangements with anyone in the field except Curtiss."[145] Who was Glenn Curtiss? He was an inventor,[146] manufacturer and pioneer, setting flying records in craft that he both designed and manufactured. His success in the aviation field made him the unwitting target of the unrelenting animosity of the Wright brothers (particularly Orville Wright). Orville's refusals to respond to Curtiss' repeated requests to work out a deal,[147] continued after the courts upheld the Wrights' patent in 1914, until finally the government had to step in and broker the creation of the Manufacturers Aircraft Association in 1917.[148] This created a pool of cross-licensed aviation patents; patent owners were paid out of this pool according to the relative value of their patents. The impact on the Wrights' bottom line of the refusal to license Curtiss cannot be overstated. Not only did Curtiss have the largest and most profitable manufacturing corporation by 1917, his design elements (some patented and some not) were far superior to the Wrights' technology and quickly became standards in the aviation industry. Even as early as 1914, and even in the midst of a court-ordered injunction in the Wright-Curtiss patent dispute, Curtiss was receiving orders for his originally-designed aircraft from all over the world. In contrast, Orville did not make a substantial amount of money from his pioneering wing-warping patent (from either licensing or manufacturing) until he sold the company in 1915.

But perhaps the infamous Wright-Curtiss scenario was an exceptional case, the rare result of a highly intense feud between individuals who knew each other personally, between inventors who vied for the laurels of public glory even more than the reward of financial return. Indeed, there is some support for this view. Once Orville Wright sold his company in 1915 to businessmen who "were more concerned with licensing and manufacturing"[149] than Orville, the Wright company moved ahead with the business of making airplanes and making deals at a brisk pace; first merging with the successful Glenn Martin company[150] in 1916 and much later (in 1929) being merged with

---

[141] W&O, 166, 189

[142] W&O at 202 (They were grievously incorrect about this – it took competitors all of five months to reach their own solution.)

[143] By 1910, they had licensed various manufacturers around the world. Shulman at 180.

[144] It should be noted that Wilbur died in 1912.

[145] Shulman at 40.

[146] Shulman at 29; 500 inventions

[147] [Insert quote from letter… let's save ourselves]

[148] W&O, 405.

[149] Wilbur and Orville at 404.

[150] Glenn Martin sold the Army its first tractor planes, sharply reducing the death rate of pilots.

none other than the Curtiss company itself.[151] If the cool head of the savvy businessman really *is* the remedy for the irrationality of the eccentric innovator, then we would expect reasonably-priced permissions to be granted to the average consumer (i.e., user of an asset) with little impeding friction. Unfortunately, it is difficult to gather information about what happens when consumers request permission to breach use restrictions, since most consumers simply breach *without* asking for permission. Complements of the digital age, however, there is at least one unusually well-documented case that we can examine: let's call it 'the strange case of Laura Flores'. Laura purchased a legal copy of a pattern-making program called Livingsoft's Dress Shop 5 Pro, for which she paid 400 US dollars. A year later, she was tired of using the program because it was awkward to use, so she decided to sell it. She listed it on eBay and got a 100 dollar bid for it. Before the sale could be concluded, Laura and her bidder received anonymous emails warning them that the transaction was illegal, to which Laura angrily retorted that she was selling the software because it 'sucks'. The company then contacted Laura, first telling her that she could not sell her copy of the software, that the End User License Agreement (EULA) expressly forbade the transaction (there was indeed a resale/transfer prohibition in the agreement). A short while later, perhaps worried about the impact of the first sale doctrine, it backed up a bit, saying that she could sell her copy of the software, but not the key codes that came with it. Unfortunately, these codes enabled users to actually print out the patterns, so the software was useless without them. Laura was shocked, and even more shocked when she saw that another person had sold her Dress Shop Pro on eBay. Apparently, the company had given that lady permission to sell her copy. So Laura emailed the company, mentioning the other seller and asking if she could have permission to sell hers as well. Livingsoft replied that they gave permissions only when the user had "a financial or physical necessity to give up use of Dress Shop"[152], and when they requested permission "courteously, prior to attempting to sell their software." Laura did not meet these requirements: "when the president of Livingsoft contacted you personally to remind you of company policy in this matter, you informed him that your reason for selling was that the software 'sucks' ... that reason doesn't qualify for an exception to our general policy."

Although this refusal of permission did not adversely impact user innovation (or perhaps it did, since it made it difficult for Laura to buy a new pattern-making program), it is useful because it provides a concrete example of what can happen when users cannot simply breach use restrictions without detection. It also provides a good example of how a set of technological developments (the Internet and eBay), not directly designed to control circumvention, can indirectly enable the monitoring and enforcement of use restrictions. Even if Laura had refused to take down the bid, Livingsoft might have been able to convince eBay to take it down. And in this case, the point was moot because they also sent emails to the prospective buyer, who was scared into refusing to go through with the transaction, either online or 'offline'. This series of events seems difficult to square with any sort of economic rationale on Livingsoft's part. They did allow resale in certain circumstances, so this was not a case in which it could plausibly be asserted that the software would not be produced at all if there were no restriction on customer resale. Of course, requiring users to request permission for resale could represent an attempt to control the sheer number of resales, but it does not seem to be an effective way of accomplishing this. After all, a customer could simply ask 'nicely', and invent a reason which fit within the justifications enumerated in the license agreement. Livingsoft's attempts to restrict resale *may* reflect a concern about the impact that resale would have on the reputation of their software or their company (after all, why would someone resell if the product was satisfactory). However, such concerns would seem to be unwarranted. Individuals often

[151] Shulman at 228. Curtiss-Wright Corporation

[152] One example of such a reason that they gave was: "an injury preventing further sewing if they asked." Ed Foster.

resell the products that they buy, and there is no reason to think that buyers of used goods draw a negative inference from this. In addition, allowing individuals to resell would itself seem to take care of such goodwill and reputational issues. It is unlikely that an individual would publicly denigrate a product that she wishes to resell,[153] as this would adversely impact both the price that she is likely to receive, and her chances of making a sale at all. Both reputational and goodwill issues would therefore seem to point towards *allowing* the customer resale of lawfully purchased products. If economic motives cannot explain Livingsoft's actions, then what could have motivated them to act the way they did? It seems clear that plain old human emotions were the culprit here. In short, Livingsoft's president was angry at Laura Flores: first, for not asking in a 'courteous manner' for permission to sell her software; and second, for telling him that his company's product 'sucked'. And this anger motivated him to deny her the permission to resell which he had granted to the other eBay seller. And this was no lowly employee, it was the president of the company, so this is hardly a question of an agent (employee) acting against the interests of his principal (employer). What lessons can we take from this example? Surprising as it may seem, corporate executives are human beings too, and as such, they suffer from the same motivational defects that encumber and ensnare the rest of us. Therefore, their propensity to grant permissions purely on the basis of good business sense cannot be presumed.

iii) Incumbent Myopia:

[This section is in progress.]

iv) Endowment Effects:

In the first three categories, value-enhancing permissions were denied or otherwise hindered because preferences did not respond, for a variety of reasons, to price. Here, the situation is somewhat different: preferences respond to price, but the price of a permission varies with an individual's ownership of it. The 'endowment effect' refers to the phenomenon that occurs when the price that we are willing to pay for an asset that we *do not* own is less than the price that we demand for an asset that we *do* own[154]. According to traditional economic theory, there should be no difference between these two amounts: if you are willing to pay 10 dollars for an item, you should be willing to sell it for 10 dollars. But the presence of an endowment effect means that although you may be willing to pay 10 dollars for an item, once you own it, you may only be willing to sell it for 12 dollars. This may hinder the movement of a resource to its higher-value user. Suppose, for example, that A values an item at 10 dollars and B values the same item at 11 dollars. A happens to be first in line at the store, and he buys the item for 10 dollars. B says to A: 'Hey, I'll buy that from you for 11 dollars.' Although A only valued the item at 10 dollars before he purchased it, because he now owns it, the endowment effect kicks in, and he will only accept an amount greater than 10 dollars. If A is now only willing to sell at a price of 12 dollars,

[153] Recall that Ms. Flores only stated that the 'software sucks' in a private email, and only after she had received anonymous emails (which she later discovered were from the company) warning her that it was illegal to sell. It is highly unlikely that she would make the same statements to a prospective buyer, especially one who had bid one hundred dollars for the product.

[154] The existence of endowment effect has been hotly debated. That debate will not be examined here, but I will insert citation evidence on both sides.

then B, the higher-value user, will not now be able to obtain the resource, even though if he and A had bid side-by-side, the item would have be purchased by B.

A similar effect might interfere with the granting of a permission to reuse, resell or modify an asset: even if the owner of the use right places a lower value on the use right than the asset user, the endowment effect might result in a refusal of permission. Endowment effects are distinct from the type of 'personhood' effects that have been examined by Radin, which tend to attach only to irreplaceable objects, such as wedding rings, to which we develop emotional attachment. In contrast, endowment effects have been shown to attach to even simple, fungible assets like coffee mugs. And although endowment effects have been shown to be more prevalent in the case of tangible objects (e.g., experimental subjects displayed endowment effects in the case of coffee mugs, but not vouchers for coffee mugs), the effect is also displayed in 'inherently' tangible assets, like stock options. [This section will be expanded.]

IV: Drawing the Line Between Governance and Exclusion: The Need for Search

Not surprisingly, just as each theory of the *numerus clausus* has its own conception of what the *numerus clausus* is, each also has a unique prescription for the robustness of the principle. In the 'third party information costs' account, the measurement and error costs that occur when new property forms are allowed, are balanced against the frustration costs generated when new forms are prevented. In the 'verification of divided rights' theory, the costs of verifying divided rights are balanced against the value of creating those divided rights. So where should the law draw the line between governance and exclusion? In this theory, the primary function of exclusionary rules is to allow users to search for new uses of assets. Accordingly, the need for the *numerus clausus* should turn on the need for search in various circumstances, and for various asset types. The need (or lack thereof) for search can be divided into four broad categories: 1) search is desirable, and users have advantages in undertaking that search; 2) search is desirable, but non-users should coordinate that search; 3) search is undesirable; and 4) search requires flexibility in asset characteristics. In the first category, exclusionary rules preserve the ability of users to seek out new uses of assets, and so the *numerus clausus* should be robustly applied to asset types that fall within this category. In the second category, special asset characteristics dictate that non-users should coordinate the search for new uses, warranting some loosening of the *numerus clausus*. In the third category, the search for new uses of assets is undesirable, for a variety of reasons which we will specify in some detail below. Because search is not needed in these cases, novel governance rules are generally unobjectionable, and the *numerus clausus* is at its weakest in this category. In the fourth category, which essentially encompasses financial instruments and trusts, the 'search' is not for new uses of tangible assets, but rather the search for new methods of financing. Thus, flexibility in the characteristics of the asset itself (whether it is a cash instrument or a derivative instrument) is required to create these new arrangements. We will discuss each of these categories in turn.

i) Search is Desirable and Users Have an Advantage in Undertaking that Search:

Quite simply, where there is a need to search out new uses of assets, then there will be benefits to allowing the possessor of the asset to operate under a rule of exclusion with respect to that asset. We have already examined (in Part 2) several case studies of user innovation, and

analyzed the various reasons why users will have advantages in the problem finding and the problem solving that are generated by asset use. We will not review them here. In general, there is far more need to search out new uses in the case of chattels and intellectual property: finding new uses of resources, making asset modifications, and creating interoperable products is critical to ongoing innovation in these areas. In many cases, these innovations will be revealed and will be adopted by fellow user-innovators and even manufacturers. In these cases, enforcing use restrictions will negatively impact innovation across a large number of both users and producers. In other cases, there is simply no market for the results of user innovation. For example, some modifications might be so unique to each user that there is simply no way that anyone else could ever discover the need for, or actually make, the modification. In other cases, it may simply be cheaper for the user to implement the modification himself. In such cases, there is simply no good reason to enforce a use restriction. [This section is in progress.]

ii) Search is Desirable, but Non-Users Should Coordinate That Search:

[This section is in progress. It essentially encompasses IP. IP appears in both i) and ii) because the rights are created to provide incentive, but they must also be limited to allow sequential innovation.]

iii) Search is Undesirable:

Just as there are cases in which we want users to search for new uses of assets, there are situations in which we do *not* want such searches to occur. There are two principle reasons why we might not want new uses to be found for, or made of, an asset. First, we might want to *control* the use of an asset. Controlling the use of an asset will not be discussed in any detail in this work, since it is largely the province of criminal law or other government regulation. Consider dangerous substances or other items, such as guns or nuclear material. It is easy to see why we might want to limit or proscribe the ability of users to use, experiment with, or even possess, such items via government regulation. Note what is *not* included in this category: private control of an asset that is designed to implement a business model (e.g., licensing restrictions that limit the ability of users to modify an asset) or prevent arbitrage (e.g., licensing restrictions that prevent asset transfer, sale, lending or rental). Although some scholars have argued that such private controls are desirable, the perspective adopted in this work is that such schemes are antithetical to user innovation, and should be enforced, if at all, only as a matter of strictly interpreted contract law. The appropriate handling of such governance rules will be discussed in more detail in Part 5.

The second reason we might not want a search for new uses is that we might want to *fix* the use of an asset. Why might we wish to do this? Fixing the use of an asset (and thus restricting search) will make sense: 1) when we want to facilitate community building; and 2) when we want to facilitate the expression of non-market values. [This section is in progress.]

Although land servitudes and conservation easements both involved efforts on the part of landowners to control the use of land after they parted with possession, these efforts generated very different institutional responses. Why did the judiciary respond to land servitudes by loosening the strictures of the *numerus clausus*, but remain stubbornly resistant to conservation

easements, thus necessitating a legislative solution? Perhaps if the English had not loved their gardens with such ardor, the judiciary might have been more reluctant to abrogate the absoluteness of the property grant (and this reluctance might, in turn, have been imported into the jurisprudence of the American colonies). But even without the enticement of the English rose, it is likely that land servitudes were, quite simply, somewhat inevitable. An island nation would be keenly aware of the necessity of coordinating land usages, and a genteel society would be sensitive to the need to create appropriately refined communities. And even in the wild expanse of the American colonies, the high rhetoric of absolute property ownership, so crucial to the fomenting of a revolution, would be expected to give way as the realities of urban living made themselves known. But why was the judiciary, so creative in responding to the need for building communities, so resistant to private conservation efforts? It might simply be a story of timing and lag – perhaps if given more time, judges might have eventually responded to changing social norms, carving out new exceptions to the *numerus clausus* that were capable of handling the novel characteristics of conservation easements. Perhaps conservationists were simply too impatient to wait for the careful and cautious wheels of judicial innovation to turn. Or, it might simply be a case of a legal rule taking on a life of its own, creating a kind of judicial blindness in which the restrictions on the creation of servitudes became an end in themselves (i.e., which had to be followed), rather than a means to an end (i.e., which could be altered according to changing circumstances). But there is another way to look at the differential institutional resolution of these cases, one which is in harmony with the underlying ethos of property ownership. Servitudes that touch and concern the land, and require a dominant and servient tenement (for example), facilitate efficient land development by allowing adjacent land owners to coordinate their activities. They thus serve the primary goals of promoting land development, and fostering community harmony. In contrast, historical and environmental preservation easements place the goals of land and community development *below* those of protecting the environment and preserving the past. Conservation schemes *hinder* property development in order to protect these competing interests, and this difference provides the key to understanding the appropriate institutional response. Perhaps the removal of real estate from the normal course of development is so fundamentally antithetical to the core goals of property ownership that it should *not* be expressed at the whim of private individuals, but only when given the imprimatur of the entire polity. And so, when it comes to environmental and historical preservation schemes, perhaps the social judgment of the legislature was not only required as a matter of expediency, but was eminently appropriate as a matter of policy.

iv) Search Requires Flexibility in Asset Characteristics:

Up to this point we have been discussing tangible assets that have an intangible (use) dimension, such as land and chattels, and intangible assets, such as copyrights and patents, that have a tangible dimension (the copy or embodiment). We now move on to a category of assets which is wholly intangible in its relevant parameters: financial instruments. [This section is in progress.]

V: The Unappreciated Importance of a Structural Background of Exclusionary Rules

The law, particularly the law of property and its various ‘doctrines’, is often criticized as being too amorphous, too fact sensitive, too unpredictable, too subject to circumvention by

determined parties. These criticisms have been aimed at everything from the 'fair use' doctrine in copyright law, to the entire law of servitudes. It might similarly be objected that the *numerus clausus* is far too nebulous to be anything other than an infinitely malleable *deus ex machina*, which can brought down from the legal heavens whenever judges' guts tell them that things 'just aren't right'. There is some truth to this view – certainly the courts often spend more time asserting 'the absurdity' of novel forms of property, than searching for a solid justification for their gastric distress. Indeed, the argument has been made that the *numerus clausus* does not exist at all, since parties can often make their novel arrangements look and act like the old, familiar ones through a bit of creative lawyering.

The implication of this is, of course, that we should therefore abandon the 'legal fiction' entirely, and open the floodgates to the ingenuity of private parties, who simply want to enforce their arrangements. But what if we were to look at it another way, as if the vagueness and flexibility were part of the design,[155] part of what makes the economic system 'work'? It is easy to see how flexibility might be a useful feature of the law: new economic situations arise every day, and we want a legal system that can deal with them. But, that sort of flexibility is not what I am trying to get at here, since the principle we are concerned with is criticized for its *lack* of flexibility. So perhaps vagueness is indeed the better term: vagueness in terms of lack of definite boundaries. What I mean is this: perhaps it is enough that the law maintains a structural background of rules of exclusion, even though the optimal level of exclusion can never be identified.[156] On the one hand, governance rules are a prime source of transaction costs (because they require that permissions flow from, and payments flow to, previous owners), and rules of exclusion reduce these costs (because by definition, they free downstream owners from these costs). On the other hand, coordination, which is the hallmark of governance rules,[157] is also required for economic activity to occur. Indeed, much of economic analysis has concerned itself with the conditions under which people will abide by their agreements, despite the temptation for rational actors to 'defect' after part performance whenever they would be better off in doing so. Given the importance of both competition and coordination to economic performance, the assertion here is that it is sufficient (and necessary) if exclusionary rules are plentiful enough that they form a structural backdrop within which governance rules be organized. Or, in other words, it is both sufficient and necessary to economic performance that governance rules, and their attendant transaction costs of permissions and payments, represent only a certain percentage of economic activity. I am heartened in this assertion by some similar sentiments expressed by Cheung in his study of economic activity in communist China:

> "[T]ransaction costs *as a proportion of gains from specialization* provides a critical measure which… very largely explains observed differences in the wealth of nations…*if this ratio is reduced just a little, a significant increase in wealth would follow*… . My favorite example is China under communism… under the communist regime transaction costs as a proportion of gains from specialization was very large. This is why the Chinese were so poor in their communist days."[158] (Emphasis added.)

We know that transaction costs will never disappear entirely - the world of zero transaction costs is truly a theoretical fiction. But the conundrum for economics posed by transaction costs goes deeper than this, for transactions costs have a 'dual nature'. On the one hand, transaction costs are a

---

[155] I do not mean to imply that the judges intended this 'design', merely that the overall scheme happens to work, perhaps by virtue of the fact that no one set out to design it that way.
[156] Although Smith has made such an attempt, his conclusions are not adopted here.
[157] Even though this 'coordination' is often non-negotiated, and thus unilaterally imposed. For this reason, I avoid using the term 'cooperation'.
[158] The Transaction Costs Paradigm at 517.

drain on the economy, and an impediment to economic activity, for they can literally raise the cost of a transaction so much that an otherwise beneficial rearrangement of rights is simply 'not worth it'. But they are also necessary to economic activity, and they comprise a sizable amount of GNP in every well-functioning, modern economy. But it is difficult, if not impossible, to precisely identify where the 'unnecessary' transaction costs (those we should seek to eradicate) end, and the 'necessary' transaction costs (those that are necessary to a functioning market) begin. This difficulty is illustrated by the controversy in economics over what a transaction cost actually *is* (definitions have included everything from the costs of transportation, to the entire activity of the service sector), and the related problem of how to measure them. This difficulty arises in economics because it is inherently difficult (if not impossible) to cleanly separate 'transactions' from 'production'. But this challenge is not unique to economics: it is the very same problem that the law has in drawing the line between governance and exclusion. A familiar example from patent law may help to illustrate the point. On the one hand, we do not want to force every inventor to 'work' his patent. He may be a brilliant inventor, but an ineffective manufacturer or businessman. Owners of patents (and other intellectual property) are thus often licensors and managers of portfolios, rather than producers themselves. In general, licensing a portfolio of intellectual property assets is seen as an acceptable facet of industrial organization, particularly in the 'post-Chandlerian' era where production is ever more disaggregated - where the central direction of the firm is replaced by an increasing number of input providers, linked by a complex series of licensing arrangements. On the other hand, something starts to look amiss (at least in the opinion of some commentators) when an individual or entity purchases a number of valuable patents, waits until others have developed a product which infringes the patent, and then swoops in to demand steep licensing revenues under threat of injunction. Instead of entrepreneurs, such individuals have been called 'patent trolls' – and they are seen by some as opportunistic leeches, burdens on the system of productive activity.

Despite the moral conviction with which these 'trolls' are criticized in the scholarly literature, business community and popular press, it is not immediately clear why their activity is singled out for such approbation. If anything, the only difference between the case in which an entity (be it the inventor or the patent owner) licenses his patent before manufacture takes place, and the case in which he demands payment after manufacture has already begun, is the likely division of profits. A manufacturer may be less willing to walk away from negotiations once expensive investments have already been made, and so may pay more for the right to continue manufacture, than he would have paid for permission to begin manufacture. The oddity of the reaction to patent 'trolls' becomes particularly apparent when we consider that the hallmark of a property right is the right to exclude - a right which is not taken away merely because the owner refuses to maximize the value of the resource. There is thus an inherent conflict between the economic incentive provided by a property right, and the risk that a particular owner will either refuse to maximize its value, or will raise the cost of productive activity for others. Like all institutions, then, property reduces or eliminates some costs, but also introduces new ones. It is these costs – the costs that property ownership inevitably imposes on sequential innovation - that the legal preference for rules of exclusion, and its robust suspicion of governance rules, seeks to mitigate. By allowing downstream owners to freely breach governance rules, rules of exclusion both lower the cost, and increase the incidence, of future innovative activity. But there is something very special about how the law achieves this result. The law of property reduces transaction costs, not by facilitating the 'costless' implementation of governance rules, but by *eliminating the need for a transaction completely*. The costs of some transactions (those which seek to implement governance rules after assets have permanently changed hands) are thus raised, while the costs of other activities (those involving new uses of assets) are reduced. This heavy-handed interference does not happen just once, of course, but every time an asset is transferred to a new 'owner'. In this way, property law has an iterative, dynamic aspect built deep into its design.

The identification of governance rules as a fertile source of transactions costs, and the consequent breakdown of transaction costs into elements of payments and permissions, provides us with a rough way to differentiate between 'necessary' and 'unnecessary' transaction costs. If you are simply free riding on the human capital of transferees, if you are merely adding a layer of costly payments and permissions without adding any independent value,[159] then there is little reason for the law to enforce such an arrangement beyond the bilateral case (which at least has the force of actual agreement behind it). Usually, the behavior that we might call 'free riding on the future' is identified with government officials, but as the 'patent trolls' example suggests, a private party can be no different from a government bureaucrat in this regard. Whether we are bribing government officials for permission, or paying private parties for that permission, the result is the same: the cost of productive activity is raised, if not blocked altogether. It is often asserted that such blockages are a concern vis-à-vis government actors but not private actors, for two reasons. First, government officials allow inefficient political motivations to guide their decisions, while private parties are guided only by the rationality of the profit motive. Second, it is said that government officials cannot be held to their bargains, while private parties can be, and this has a disciplining effect on the capricious behavior of private parties. While these objections have some merit, they capture only part of the reality, particularly when we consider that private parties can suffer from cognitive and motivational idiosyncrasies of their own – characteristics which can make them act as 'unreasonably' and autocratically as a bureaucrat. A private actor may indeed be better than a government bureaucrat, and certainly it is accepted that this is most often true. But over time, the independent wills of *many* private actors will produce better results in the search for new uses of assets than the restrictive wills of only a few individuals. And so, spreading property rights in assets over multiple persons, by forcing the transfer of restriction-free bundles, at least partially offsets the deleterious impact that the idiosyncratic 'whims' of private parties can impose on future owners.

Of course, property ownership is nothing other than the right to impose private whims upon the world, and so it is inherently difficult to draw the line between governance and exclusion with a steady hand.[160] This is why tangibility will always play an important role in the law of property. In reality, there is really no way to differentiate between 'governance rules that raise transaction costs' and the normal incidents of property ownership. The power to contract over an asset is, after all, one of the most important sticks in our bundle of property rights. This is why the law must use the device of tangibility to do the work (i.e., full use travels with permanent possession). Tangibility is, quite simply, the best we can do to separate the exclusionary sheep from the governance goats. And because tangibility operates to unify full use with possession *only* when the physical transfer of an asset has achieved a certain degree of permanence, the coherence of the concept of property is maintained. Tangibility (and its offspring, possession) must walk a fine edge in property law,

---

[159] Some might be concerned that this is merely a 'choice of contracts' issue. That is to say, choosing between employees, sharecroppers and piece workers, for example, is a simply a choice between different contractual arrangements – none of which transfer the residual - so perhaps licensing restrictions are simply another point along this continuum. However, we need only ask ourselves whether every transferee of an asset should automatically be considered to be bound by restrictions imposed by the transferor. In arrangements involving employees, for example, there are other factors that compensate the transferee for his lack of asset ownership. Though it might sound fantastical to equate the situation of a buyer with the situation of an employee or servant, the equivalence is inherent in the assertion that all purchasers should be bound to property restrictions because they implement the business model of the seller (for example, maintaining a certain number of sales, at certain prices, by preventing purchasers from engaging in arbitrage). In other words, should consumers be forced to be part of the firm? For more discussion of this, see Part 5.

[160] Incidentally, this need for rough line drawing applies to both rights as between private persons, and rights as between private persons and the government. At some point, government zoning can restrict property use to such a degree that it literally dispossesses the property owner of his property. At this point, an expropriation has occurred. See, for example, Pennsylvania Coal. In the modern regulatory state, the ability of a property owner to insulate his property from government regulation has been greatly attenuated.

balancing itself between opposing chasms of incoherence. On the one hand, if possession is required at all times for property ownership, there is no property ownership at all. After all, to have an enforceable property right while not in physical possession of an asset is the hallmark of property ownership.[161] Possession may be 'nine-tenths of the law', but if possession is required at all times, then 'ownership' literally has no meaning. Think about what would happen to the concept of property in a world where adverse possession vested instantaneously upon every physical transfer, for example.

On the other hand, if the possession of an asset is *completely* separated from ownership of that asset, the concept of property once again falls by the wayside. Consider a world in which every person along the chain of transfer of an asset were subject to the claims of former owners. In other words, we could contemplate an ongoing resale/reuse royalty structure for all property, not simply intellectual property. This is not as fantastic as it seems. Consider the following hypothetical: I own a garden variety plot of land, and for 20 years I take care of it. I finally sell my average land for an average price, and the next day it is discovered that my garden variety plot of land is actually very special. It is the only place where a plant that cures cancer can grow. Furthermore, the land can only support this miraculous plant because I took such good care of it for 20 years. The 'owner of a day' sells the land on his second day of ownership for 1 billion dollars. Might it not be fair for me to receive a resale royalty under these circumstances? We could also imagine a more extreme (and perhaps less sympathetic) version of the resale/reuse royalty. Suppose the legal system adopted a severe version of 'first in time is first in right' principle: once an initial property owner is identified, ownership never transfers, even after the possession of the asset changes permanently. In this scheme, all uses made by all future transferees must be paid for, and all are subject to the permission of the 'initial' owner. Such a 'pay-per-use' or 'permission' society looks more like a web of feudal relationships, an extended scheme of bailment in which possessors of assets not only temporarily hold them for the benefit of an other, but permanently *use* them for the benefit of another. In other words, for the concept of property to have any meaning at all, we at all times need islands of exclusion in a sea of governance. Conveniently, the law rarely has to even admit the possibility of such scenarios (let alone their desirability), because it simply falls back on the device of tangibility. Thus, possession has important work to do, not just on first possession, but whenever an asset is permanently transferred.

Of course, it is impossible to say at any particular point in time whether the law has drawn the line between governance and exclusion at the 'right' place, whether the *numerus clausus* as a whole is 'too open', or 'too closed'. What we can say, however, is that a property system with a *numerus clausus* imposes fewer transaction costs on sequential innovation than one without. The fact that this provides us only with a relative ranking is not fatal to the usefulness of the observation, as we do not need to be able to measure transaction costs to say something useful about them, we need only be able to rank them.[162] Of course, it is difficult to develop testable implications (precise, predictable results in individual cases) from a general theory, especially when that theory seeks to account for something as complex and varied as the *numerus clausus*. The inability of legal theory to generate testable implications has been repeatedly criticized by economists. In fact, this criticism has recently been aimed at the previous theories of the *numerus clausus*, as well as the previous theories of servitudes. This disparity between legal and economic theory is not surprising. The law is quite comfortable with general theories; economics is not. In part this stems from the distinct natures of the enterprises: economics self-consciously strives for the precision of a 'science'; while the law stubbornly resists absolutes. It is true, for example, that as recordation schemes developed, the *numerus clausus* was somewhat weakened. But is also true that the *numerus clausus* continues to

[161] Hume?
[162] Cheung.

persist despite improved means of providing notice. Similarly, it is true that the law does not put an absolute prohibition on the creation of new property forms, so the number is not truly ‘closed’. But it is also true that the courts spill a great deal of judicial ink on analyzing, crafting, and justifying the rules that restrict the ability of private parties to either create new categories or gain entry into existing ones. But perhaps this lack of absolutes is, in the end, the real strength of the law, rather than its weakness. Perhaps it is the only reasonable way to deal with the endless variety inherent in human beings and human activity, both of which defy easy prediction or categorization. But general theories would seem to have their place in economics too, or else it risks saving the idiosyncratic trees while the forest burns. We turn now to have a look at a few of those trees, and the general state of the forest, in Part 4.