A User Innovation Theory of the *Numerus Clausus*

By

Leah Marie Theriault

A thesis submitted in conformity with the requirements
for the degree of S.J.D.
Graduate Faculty of Law
University of Toronto

© Copyright by Leah Marie Theriault, 2011
A User Innovation Theory of the *Numerus Clausus*

Doctor of Juridical Science (S.J.D.) 2011

Leah Marie Theriault

Faculty of Law
University of Toronto

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, built upon 1) the pivotal role that rules of exclusion play in fostering user innovation, and 2) the role that psychological ownership plays in preventing recontracting around governance rules, 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 (governance), replacing them with rules of exclusion (exclusion). Exclusion unifies rights of use and possession in assets, while governance separates, to a greater or lesser degree, possession from use rights. Full user, sale and the policy against restraints on alienation are the paradigmatic examples of exclusion; while governance is exemplified by servitudes and contractually-burdened assets. Exclusion plays a critical role in user innovation because it allows the possessors of assets to unilaterally seek out new uses of those assets. Whenever the law replaces governance with exclusion, user innovation is more likely to occur because the 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. In addition to negatively impacting user innovation, governance hinders recontracting because both possession and legal entitlements (rights of use in an asset) give rise to feelings of psychological ownership, and individuals will not recontract over uses that they feel they already ‘own’. The user innovation 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 (where users generate a significant amount of innovation), and why it has been somewhat attenuated in the area of real property (where we restrict search in order to facilitate coordination of land uses). It also explains why the law enforces the principle even when the cost of providing notice of governance rules is low.
Acknowledgments

An S.J.D. thesis is not a joint work, but it is a joint effort. And so, there are far more people to thank than there are words adequate to thank them. To my supervisor, Bruce Chapman, whose clarity of thought and keenness of mind have made me a better scholar, and who has read through more discarded verbiage than perhaps any other supervisor in the history of the faculty. To Abraham Drassinower, who stayed on through my move from books and movies, to skateboards and baking soda. To Michael Trebilcock and Ariel Katz, who made it possible for me to finish, not just in time, but in time for my birthday. And to Peggy Radin, whose work was there at the beginning, and who was there in person at the end. You have all greatly improved not just this work, but the work that I will do in the future.

To Julia Hall, who deserves roses every day. To the faculty, staff and students at the University of Toronto and Washington University in St. Louis, for showing me what an educational institution can and should be. To my family, who helped in so many ways over these long years. And finally, to Chuck McManis, who helped restore my faith, and to my Mother, who helped me cross the finish line. Without all of you, doing things the hard way would not have been possible, and the hard way is the only way I know.

Together, you are all living proof of one simple fact:

The only debts worth incurring are the debts you can never repay, from people who will never ask you to repay them.
PART 3  DIVIDED RIGHTS VERSUS FRAGMENTED RIGHTS .......................................................... 123

PART 4  PSYCHOLOGICAL OWNERSHIP VERSUS ANTICOMMONS THEORY .......................... 125

INTRODUCTION .......................................................................................................................... 134

CHAPTER 4: THE USER INNOVATION THEORY AND TANGIBLE PROPERTY .................... 133

INTRODUCTION .......................................................................................................................... 134

PART 1  PROMOTING USER INNOVATION .............................................................................. 135

1) The Legal Policy Against the Creation of Novel Servitudes .................................................. 136
   i) Land ........................................................................................................................................ 137
   ii) Chattels .................................................................................................................................. 139
   iii) Scholarly Criticism of the Legal Policy Against Servitudes .................................................. 142
   iv) Restrictions on the Creation of Novel Servitudes Promote User Innovation ......................... 143

2) The Law Has a Greater Tolerance for Servitudes on Land Than for Servitudes on Chattels .... 145
   i) Search is Desirable and Users Have an Advantage in Undertaking that Search ................. 147
   ii) Search is Undesirable ...................................................................................................... 149
      a) Coordinating Land Uses .......................................................................................... 149
      b) Expressing Non-Market Values ................................................................................. 151

PART 2  VALIDATING PSYCHOLOGICAL OWNERSHIP ....................................................... 153

1) Divided Rights Versus Fragmented Rights ............................................................................ 154

2) The Common Law Does not Recognize Permanent Leases ............................................... 155
   Possession and the Duty to Return .................................................................................... 156

3) Notice Does not Determine the Boundaries of the Numerus Clausus .................................... 158
   A) The Legal Approach to Notice .................................................................................. 159
      i) ‘Bare Notice’ Chattel Cases Remain Anomalous ......................................................... 160
      ii) IP Copies and Notice; The Realities of the Transaction ............................................. 163
   B) The Scholarly View of Notice .................................................................................... 166
   C) The User Innovation Theory Accounts for the Legal Approach to Notice ....................... 168
      i) From Information Asymmetries to Asymmetries in Cognition and Motivation ............. 168
      ii) Psychological Ownership, Notice and the Bargaining Process ................................. 170
      iii) User Innovation and the Mechanisms of Governance ............................................ 172

CONCLUSION ............................................................................................................................ 175

CHAPTER 5: THE USER INNOVATION THEORY, INTELLECTUAL PROPERTY, AND TECHNOLOGY .. 176

INTRODUCTION .......................................................................................................................... 177

PART 1  THE USER INNOVATION THEORY AND INTELLECTUAL PROPERTY LAW ........... 177

1) Search is Desirable, but Non-Users Should Control the Search ............................................. 181

2) The User Innovation Theory and the Structure of Intellectual Property Law ...................... 181
   A) Intellectual Property Works are Like Land ............................................................... 181
   B) Copies are Like Chattels ............................................................................................ 186
      i) First Sale .................................................................................................................. 187
      ii) Criticisms of First Sale/Restraints on Alienation ................................................... 190
      iii) The ‘Mere’ Right to Rent/Resell Fosters User Innovation ....................................... 193
      iv) The Future of First Sale ......................................................................................... 198

PART 2  THE FUTURE OF THE NUMERUS CLAUSUS - THE CHALLENGE OF TECHNOLOGY.. 200

1) Digital Rights Management and the Digital Millennium Copyright Act ............................. 205

2) Do Technological Controls Interfere with the Development of Feelings of Psychological Ownership? .......................................................... 210

3) The Legal Response to Technology and Business Alternatives to Use Restrictions ............ 213

CONCLUSION ............................................................................................................................ 217

CONCLUSION ............................................................................................................................ 219

THE UNAPPRECIATED IMPORTANCE OF A STRUCTURAL BACKGROUND OF EXCLUSIONARY RULES .... 223

BIBLIOGRAPHY .......................................................................................................................... 235
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 any owner.”¹

Why are property rights, in the eyes of the law, so much less customizable than contract 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 limitations on the customizability of property rights. Indeed, economic considerations generally support the imposition of land use restrictions, for example, to encourage property owners to make the investments necessary for the efficient development and use of resources. 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 it is best understood as a convenient way to refer to the ‘Holmesian impulse’ in property law: “It would be most unfortunate and unexpected if it should be discovered at this late day that it was possible to impose such a

---

2 See e.g. Carol Rose, Servitudes, Security and Assent: Some Comments on Professors French and Reichman, 55 S. CAL. L. REV. 1403, 1403 (1982): “[W]e tolerate these “dead hand” arrangements because they provide a long lasting security for land development and encourage property owners to invest in the long term improvements that are essential to the productive use of real estate. As Jeremy Bentham told us over a century ago, secure expectations of return are a sine qua non of enterprise; if the community wants to encourage my neighbor to invest time and effort in a solar heating panel, it had better let him agree with me that I and my successors will keep my trees from overshadowing his roof.” (Citations removed.)

3 See e.g. Glen Robinson, Personal Property Servitudes, 71 U. CHI. L. REV. 1449, 1462 (2004) [Robinson]: “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.”

4 The law considers the creation of a new kind of servitude to be a ‘new’ type of property right. Economists, on the other hand, generally consider servitudes to be a different division of rights.
qualification upon a fee.” 5 This impulse is ubiquitous throughout property law. 6 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. 7 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 thesis that the issue of the scope of property rights, the issue of servitudes, 8 and the ‘license or sale’ question, are at heart one issue, which is better addressed by a general theory - under the numeral clausus umbrella – than by isolated justifications or treatments.

Does all this legal interference with the will of private actors make good economic sense? Until the past decade or so, the numerus clausus, which is so central to the universe of property law, 9 received little attention from either economists or ‘law and economics’

5 Johnson v. Whiton, 159 Mass. 424, 426; 34 N.E. 542, 543 (Mass. 1893) [Whiton]. It should be noted, however, that Holmes did not always apply this solicitude to the business plan and intellectual property cases. See e.g. his dissents in Miles Medical Co. v. John D. Park & Sons Co., 220 U.S. 373, 411 (1911): “I think that, at least, it is safe to say that the most enlightened judicial policy is to let people manage their own business in their own way, unless the ground for interference is very clear.” (Arguing in support of price maintenance scheme vis-à-vis retailers,); and Motion Picture Patents Co. v. Universal Film Mfg. Co., 243 U.S. 502, 519: “I suppose that a patentee has no less property in his patented machine than any other owner… I cannot understand why he may not keep it out of use unless the licensee, or, for the matter of that, the buyer, will use some unpatented thing in connection with it.” (Arguing in support of a tied sales scheme for patented goods.)
6 See e.g. Thomas W. Merrill and Henry E. Smith, Optimal Standardization in the Law of Property: The Numerus Clausus Principle, 110 YALE L.J. 1, 3-4 (2000) (citations removed) [Merrill & Smith]: “Every common-law lawyer is schooled in the understanding that property rights exist in a fixed number of forms. The principle is acknowledged - at least by implication - in the "catalogue of estates" or "forms of ownership" familiar to anyone who has survived a first-year property course in an American law school. The principle, however, is by no means limited to estates in land and future interests; it is also reflected in other areas of property law, including landlord-tenant, easements and servitudes, and intellectual property.”
9 Even though it is largely implicit in the common law (unlike the civil law). See e.g. Merrill & Smith, supra note 6 at 4: “In the common law, the principle that property rights must conform to certain standardized forms has no name. In the civil law, which recognizes the doctrine explicitly, it is called the numerus clausus - the number is closed.”
scholars.\textsuperscript{10} At present, there are two principal economic theories\textsuperscript{11} of the \textit{numerus clausus}: 1) the third party information costs argument developed by Thomas Merrill and Henry Smith;\textsuperscript{12} and 2) the costs of verifying divided rights rationale put forward by Henry Hansmann and Reinier Kraakman.\textsuperscript{13} These theories are important and seminal contributions to the literature, and they have significantly advanced our understanding of the \textit{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.\textsuperscript{14} Private law, in contrast, is squarely concerned with the impact on those directly involved with the novel property form.\textsuperscript{15} Second, both theories explicitly state that as technology lessens the cost of providing notice of novel forms, the law’s restrictions on the number of forms should

\textsuperscript{10} Rudden wrote one of the first articles analyzing the doctrine in detail. See Bernard Rudden, \textit{Economic Theory v. Property Law: The Numerus Clausus Problem}, in \textit{Oxford Essays in Jurisprudence: Third Series} 239 (John Eekelaar & John Bell eds., 1987) [Rudden]. At that time he wrote: “the current literature offers no economic explanation of the \textit{numerus clausus} (that is, the limited number of allowable property forms), but seems largely to ignore its existence” at 241. Merryman was one of the first scholars to identify a \textit{numerus clausus} within the American common law, see John Henry Merryman, \textit{Policy, Autonomy, and the Numerus Clausus in Italian and American Property Law}, 12 Am. J. Comp. L. 224 (1963).

\textsuperscript{11} There is actually a third theory, the anticommons theory first put forward by Michael A. Heller, \textit{The Tragedy of the Anticommons: Property in the Transition from Marx to Markets}, 111 Harv. L. Rev. 621 (1998). This theory is subject to the challenge that the law does not concern itself with \textit{fragmented rights} - i.e., the law actually allows significant fragmentation of property interests along the tangible dimension, as well as allowing multiple individuals to share ownership of assets. Thus, although the anticommons theory identifies an important set of concerns, these concerns have little to do with the \textit{numerus clausus}. For this reason, the anticommons theory will not be discussed in Chapter 1, where the principal theories of the \textit{numerus clausus} are reviewed, but in Chapter 3, where the difference between fragmented rights and \textit{divided rights} becomes important - \textit{see infra} footnote 43 and accompanying text. Although the user innovation theory of the \textit{numerus clausus} is, like Heller’s theory of the anticommons, concerned with underuse, in the user innovation theory underuse does not arise from excessive fragmentation, but from \textit{any} separation of possession from use rights. Francesco Parisi has built upon Heller’s work in formulating a theory of servitudes, and that theory (and its relation to the one offered here) will also be discussed in Chapter 3. \textit{See Chapter 3, Part 4, Psychological Ownership Versus Anticommons Theory}. There are also an increasing number of non-economic theories of the \textit{numerus clausus}, but these will not be dealt with in this work. \textit{See e.g.} Nestor M. Davidson, \textit{Standardization and Pluralism in Property Law}, 61 Vand. L. Rev. 1597 (2008), Hanoch Dagan, \textit{The Craft of Property}, 91 Cal. L. Rev. 1517 (2003), and Daphna Lewinsohn-Zamir, \textit{The Objectivity of Well-Being and the Objectives of Property Law}, 78 N.Y.U. L. Rev. 1669 (2003).

\textsuperscript{12} Merrill & Smith, \textit{supra} note 6 at 24 (citations removed).


\textsuperscript{14} See Merrill & Smith, \textit{supra} note 6 at 44: “the adoption of idiosyncratic property rights has an impact not only on the originating parties and potential successors in interest, but also on other market participants. Idiosyncratic rights create a common-pool problem, which does impose external costs on third parties”; and Hansmann & Kraakman, \textit{supra} note 13 at S374: “We agree, in particular, that third-party information costs are central to the law’s regulation of property rights.”

\textsuperscript{15} \textit{See e.g.} Hanoch Dagan, \textit{The Law and Ethics of Restitution} (Cambridge University Press 2004) at 225: “correlativity between the two parties if what distinguishes private law from regulation, whereby individuals are penalized for harms committed against society.”
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”.  

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 thesis 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. These lacunae are addressed by focusing on the role that the *numerus clausus* plays in fostering the phenomenon of user innovation; for this reason, the theory developed here is called the user innovation theory of the *numerus clausus*.

In the user innovation theory, the *numerus clausus* is a by-product of the law’s preference for rules of exclusion and concomitant antipathy to governance rules. Rules of exclusion unify all rights of use of an asset with the right of possession; whereas governance rules separate possession from use rights. This preference for the unification of use and possession gives rise to the limited menu of property forms that are characteristic of the *numerus clausus*. But the unification of use and possession is more than a pointless or misguided legal idiosyncrasy. The unification of rights of use with the right of possession is absolutely critical to the search for new uses of assets, because many innovations can be revealed only through asset use. Furthermore, possessors interact differently with assets than non-possessors (for obvious reasons), while each possessor interacts differently with an asset than every other possessor of that asset (because each individual has a unique stock of human capital). In other words, different allocations of use and possession will produce different innovative outcomes because individuals interact differently with assets, depending

---

16 Merrill & Smith, *supra* note 6 at 42 (citations removed).
17 The recent case of UMG Recordings v. Augusto, 558 F. Supp. 2d 1055 (C.D. Cal. 2008), was typical in this regard. In that case, a restriction on resale, which was clearly printed on CDs, was not enforced by the court despite the fact that all persons who handled the CDs were aware of restriction. Thus, notice was not only cheap and easy in this case (the printed label on the CD), it was effective (the reseller had both read and understood the restriction).
18 For a detailed explanation of these terms (governance and exclusion), see the discussion *infra* Chapter 1, *An Alternate Definition: A Legal Preference for Rules of Exclusion*.
19 *Id.*
20 The evidence for this proposition is discussed *infra* Chapter 2, Part 2, *Intentional Innovation*.

5
on how rights of use and possession are divided. These changes in how individuals interact with assets mean that resource values will be higher when rights of use and possession are unified.\textsuperscript{22}

Furthermore, in many cases, individuals will not contract around governance rules because individuals also transact differently over assets, depending on how rights of use and possession are divided. These changes in transacting behavior occur because both rights of use and physical possession give rise to feelings of psychological ownership, which interfere with the recontracting process.\textsuperscript{23} Thus, transaction costs will be higher when rights of use are divided from the right of possession. As a result of these changes in interacting and transacting behavior, the innovative outcomes that occur under exclusion will not be the same as those produced under governance.\textsuperscript{24}

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.\textsuperscript{25} This happens because technological restrictions on digitized content actually become part of the product that is transferred to the consumer,\textsuperscript{26} and because digital goods can provide notice of contractual restrictions on content each time the product is accessed.\textsuperscript{27}

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

\begin{thebibliography}{9}
\bibitem{22} See infra Chapter 2.
\bibitem{23} See infra Chapter 3.
\bibitem{24} Id.
\bibitem{25} See e.g. Margaret J. Radin, \textit{Regulation by Contract, Regulation by Machine}, 160 J. INSTITUTIONAL & THEORETICAL ECONOMICS 142 (2004) at 160: “The advent of Digital Rights Management Systems (DRMS’s) has the potential to read out the regulatory contouring of contract just as the advent of [technologically implemented contracts]... has the potential to read out the regulatory contouring of [intellectual] property.”
\bibitem{26} See e.g. Robinson, supra note 3; Molly Van Houweling, \textit{The New Servitudes}, 96 GEO. L.J. 885 (2008) [New Servitudes].
\bibitem{27} See e.g., Niva Elkin-Koren, \textit{Copyrights In Cyberspace - Rights Without Laws?}, 73 CHI.-KENT L. REV. 1155 (1997-1998) at 1156: “The technical ability to make any access contingent upon accepting the terms of a license allows information providers to subject all users to standard terms of use. Such contracts are often automatically enforced by the code that facilitates access to the works so that only uses that are licensed by providers become technically available to users.”
\end{thebibliography}
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 be 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.\(^{28}\)

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.”\(^{29}\) 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?\(^{30}\)

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.\(^{31}\) Indeed, the digitally-encoded restrictions themselves become ‘part of the product’, and absent some

\(^{28}\) *Id.*

\(^{29}\) Those interested in judicial eloquence might also 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.” Specht v. Netscape Communications Corp., 150 F. Supp. 2d 585, 587 (S.D.N.Y. 2001), aff’d, 306 F.3d 17 (2002). In this particular case, the court declined to find a binding contract; however, like many other courts, the court did display a willingness to recognize the new means of contractual formation that are enabled by new technologies: “Netscape argues that the mere act of downloading indicates assent. However, downloading is hardly an unambiguous indication of assent. The primary purpose of downloading is to obtain a product, not to assent to an agreement. In contrast, clicking on an icon stating "I assent" has no meaning or purpose other than to indicate such assent. Netscape's failure to require users of SmartDownload to indicate assent to its license as a precondition to downloading and using its software is fatal to its argument that a contract has been formed.” *Id.* at 595.


sort of legislation that mandates product design,\textsuperscript{32} 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.\textsuperscript{33} 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.

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 technology, such as Radio Frequency Identification tags (RFIDs),\textsuperscript{34} present the possibility that even conventional goods might some day be subject to electronic servitudes.\textsuperscript{35} Indeed, there is some evidence that producers of tangible products are already trying to use labels containing

---

\textsuperscript{32} Much in the same way that legislation regulated the ‘product’ of insurance contracts when egregious terms became the industry norm. Friedrich Kessler, \textit{Contracts of Adhesion — Some Thoughts about Freedom of Contract}, 43 COLUM. L. REV. 629 (1943).

\textsuperscript{33} See generally Robinson, \textit{supra} note 3 at 480-515.

\textsuperscript{34} See e.g. BILL GLOVER AND HIMANSHU BHATT, \textit{RFID Essentials} (O'Reilly Media 2006) at 1: “RFID stands for Radio Frequency Identification, a term that describes any system of identification wherein an electronic device that uses radio frequency or magnetic field variations is attached to an item. The two most talked-about components of an RFID system are the \textit{tag}, which is the identification device attached to the item we want to track, and the \textit{reader}, which is a device that can recognize the presence of RFID tags and read the information stored on them.” In 2007, one group of industry analysts stated that: “In 10 to 15 years RFID technology will be ubiquitous.” See V. DANIEL HUNT, ALBERT PUGLIA AND MIKE PUGLIA, \textit{RFID – A Guide to Radio Frequency Identification} (Wiley 2007) at 111. Currently, RFID’s principal areas of application are in: tracking systems (animals, individuals, vehicles, inventory); payment systems (particularly point of sale purchase); access systems (such as key cards); and control of counterfeit goods. \textit{Id.} at 1 and 73-74. More exotic applications include the use of digestible tags to track the consumption of sushi for payment purposes, see RFID in Japanese Restaurants, available at \url{http://asianewbiz.blogspot.com/2004/08/RFID-in-japanese-restaurants.html} (Last visited June 9, 2010.) The smallest RFID chip currently in commercial production, which can be embedded in a piece of paper, is 0.4x0.4 mm. See Hitachi Europe Limited, \textit{Mu Chip Data Sheet}, available at \url{http://www.hitachi-eu.com/mu/Products/Mu%20Chip%20Data%20Sheet.pdf} (Last visited June 9, 2010.) Hitachi has also developed a smaller chip, 0.05x0.05mm, which is so small that it looks like ‘powder’ to the human eye: RFID ‘Powder’ – World’s Smallest RFID Tag, available at \url{http://www.technovelgy.com/ct/Science-Fiction-News.asp?NewsNum=939} (Last visited June 9, 2010.)

\textsuperscript{35} The phenomenon of electronic servitudes is discussed in New Servitudes, \textit{supra} note 26.
restrictive license terms in order to impose restrictions on the use of their products.\textsuperscript{36} 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 \textit{de facto} and \textit{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 aside? In other words, is the legal concept of ‘sale’\textsuperscript{37} nothing more than a technological artifact?

In the interests of clarity, and because the topic of this thesis crosses so many contentious doctrinal and theoretical fault lines, a few words need to be said at the outset on what this thesis does and does \textit{not} assert. First, it does not assert, as some others have, that the common law process tends to produce efficient results.\textsuperscript{38} What this work does assert is that a \textit{structural background of exclusionary rules} is efficient.\textsuperscript{39} The fact that the law has traditionally fostered this structural background provides us with an example of an efficient legal rule, but it might be nothing more than an isolated example, rather than an inevitable by-product of the common law process. Second, and related to the first point, this work makes no statement as to relative institutional competence. Instead, it focuses on the \textit{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.\textsuperscript{40} A substantial amount

\begin{footnotesize}
\begin{enumerate}
  \item Actually, such attempts have a venerable history behind them. For convenient summaries of such attempts see \textit{e.g.} Randal C. Picker, \textit{Copyright and the DMCA: Market Locks and Technological Contracts, in ANTI TRUST, PATENTS AND COPYRIGHTS: EU AND US PERSPECTIVES} (Francois Leveque and Howard Shelanski eds., Edward Elgar 2005); and Randal C. Picker, \textit{From Edison to the Broadcast Flag: Mechanisms of Consent and Refusal and the Propertization of Copyright}, 70 U. CHI. L. REV. 281 (2003).
  \item And the \textit{numerus clausus}, and the legal preference for exclusionary rules.
  \item See \textit{e.g.} Paul Rubin, \textit{Why is the Common Law Efficient?}, 6 J. LEGAL STUDIES 51 (1977); and \textbf{RICHARD POSNER, ECONOMIC ANALYSIS OF LAW} (4\textsuperscript{th} ed. 1992).
  \item See infra Conclusion, \textit{The Unappreciated Importance of a Structural Background of Exclusionary Rules}.
  \item The perspective expressed by Williamson is typical. See \textbf{OLIVER WILLIAMSON, THE MECHANISMS OF GOVERNANCE} (Oxford University Press 1996) at 353: “Ronald Coase’s counsel that… all forms of organization are subject to failures, hence the need always and everywhere to do comparative institutional analysis, were ideas that I immediately embraced” (citation removed).
\end{enumerate}
\end{footnotesize}
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 thesis is comprised of five parts: Chapter 1 provides a new definition of the *numerus clausus* (the law’s preference for rules of exclusion); Chapter 2 contains a detailed account of how the law’s preference for exclusion promotes user innovation; Chapter 3 shows how feelings of psychological ownership interfere with the process of recontracting around governance rules; Chapter 4 takes a closer look at whether user innovation and psychological ownership can successfully account for the law’s regulation of use restrictions on tangible property; and Chapter 5 examines whether the user innovation theory can account for the structure of intellectual property law. Chapter 5 also examines the future of the *numerus clausus* more generally, in regards to both tangible and intangible assets.

Chapter 1, *The Numerus Clausus - The Law’s Preference for Exclusion*, briefly outlines previous theories of the *numerus clausus* and identifies why we need a new theory. It then goes on to equate the *numerus clausus* with a legal preference for rules of exclusion, rather than governance rules. Rules of exclusion (such as the policy against restraints on alienation) unify rights of use and rights of possession, giving the possessors of assets full use rights over the assets that they possess. Governance rules (such as servitudes) separate, to a greater or lesser degree, possession from use rights.

Chapter 2, *User Innovation - Why Exclusion is Important*, reveals why exclusion deserves the solicitude traditionally accorded to it by the law, by focusing on the phenomenon of user innovation (an approach which brings us squarely within the chain of privity of the asset). Unifying rights of use and possession plays an important role in promoting innovation because different divisions of use and possession generate different innovative outcomes. This is because individuals interact differently with assets depending on how rights of use and possession are divided. Users have advantages in innovating because a substantial number of innovations emerge only through asset use. Therefore, giving the possessors of assets full use of those assets fosters innovation (both intentional and accidental). The phenomenon of user innovation means that resource values will be higher when rights of use and possession are unified.
But the innovative impact of unifying use and possession extends far beyond the bilateral case, because users not only interact differently with assets than non-users, *each user’s interaction differs from that of every other user*. The foundation of this insight lies in the role that human capital plays in the innovative process. Human capital is the key constraint in the search for new uses of assets, and each person’s stock of it is rival, non-transferable and unique. This makes it appropriate for the law to organize property rights around the possessors of assets whenever possible (i.e., to favor exclusion over governance).

Chapter 3, *Psychological Ownership and Barriers to Recontracting*, addresses the role that effective notice should play in the relaxation of the *numerus clausus* principle. Effective notice of a governance rule (e.g., the specifics of an unusual servitude) should only result in a relaxation of the *numerus clausus* if individuals will contract around governance rules whenever it would be advantageous for them to do so. Recontracting often does not happen, however, because individuals *transact differently over assets*, depending on how rights of use and possession are divided. Both the physical possession\(^{41}\) of an asset, and the ability to refuse permission to use an asset (i.e., the ownership of use rights), foster the development of feelings of psychological ownership – feelings which go beyond the boundaries of legal ownership. This creates a situation in which both the possessor of the asset and the owner of the use rights are unwilling to recontract around divisions of possession and use. Psychological ownership represents the greatest hindrance to recontracting, but other factors – such as noneconomic motivations, shortsightedness, and the mutually exclusive choice between market and social norms - also interfere with the recontracting process.

Chapter 4, *The User Innovation Theory and Tangible Property*, takes a closer look at how the law determines when governance rules are and are not enforceable in relation to tangible property. The law draws the line between governance and exclusion along two major axes: asset type and transaction characteristics. 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 of exclusion are indicated whenever there will be benefits to searching out new uses of assets, and this will vary with asset type. It is to be

\(^{41}\) Use is also an important determinant in the development of feelings of psychological ownership. See Chapter 3, Part 1, A) 1) Psychological Ownership, Possession and Use.
expected, therefore, that the *numerus clausus* will be (and should be) more persistent in regards to personal property (where user innovation is particularly robust), and somewhat attenuated in real property (where restrictions on search facilitate the coordination of land uses\(^{42}\)). Thus, user innovation accounts for the following features of the *numerus clausus* (none of which are adequately accounted for by previous theories): 1) the legal policy against the creation of novel servitudes; and 2) the fact that the law has a greater tolerance for servitudes on land than for servitudes on chattels.

In regards to transaction characteristics, the law has a strong preference for imposing exclusion when there is no duty to return an asset, and when no bargaining process accompanies the transfer of an asset. Both of these transactions characteristics are consistent with the role that possession, and the ability to refuse permission for asset use, have in promoting feelings of psychological ownership. Psychological ownership accounts for the following features of the *numerus clausus* (none of which are adequately accounted for by previous theories): 1) the fact that the law is highly intolerant of divided rights, but highly tolerant of fragmented rights;\(^{43}\) 2) the fact that the law does not recognize permanent leases; and 3) the fact that notice alone does not determine the boundaries of the *numerus clausus*.

Chapter 5, *The User Innovation Theory, Intellectual Property, and Technology*, examines whether the user innovation theory of the *numerus clausus* can account for the structure of intellectual property law, which is the most fertile source of divided rights or use restrictions on assets. Although intellectual property law *restrains* user innovation in order to provide incentives for the production of intangible assets which are costly to produce but cheap to reproduce, intellectual property law also contains several features which are protective of user innovation, such as the fair use defense. The user innovation theory also accounts for the following structural features of intellectual property law: the idea/expression dichotomy, the nonobviousness criterion, and the first sale doctrine. Chapter 5 then looks to the future of the *numerus clausus*. In particular, it examines the special challenges that

\(^{42}\) The coordination of land uses requires that the search for new uses of land be restricted whenever the use of one parcel depends on the use of another parcel. For example, if I want to maintain my view of the harbor, I need you to refrain from building your house too high. See Chapter 4, Part 1, 2) ii) *Search is Undesirable.*

\(^{43}\) *Fragmented rights* divide rights along the tangible dimension - i.e., the law does not care if you cut a widget in half and sell half a widget. *Divided rights* separate the use of an asset from the possession of an asset – i.e., the law will usually interfere if you sell a widget subject to a use restriction. Theorists tend to equate these two categories of rights, but the law does not - and this thesis as a whole explains why. For more on the difference between divided rights and fragmented rights, see Chapter 3, Part 3, *Divided Rights Versus Fragmented Rights.*
technological developments pose to the law’s traditional means of determining the enforceability of use restrictions (in regards to both tangible and intangible property).

The Conclusion briefly reviews the argument, and discusses the general importance of a structural background of exclusionary rules to both the promotion of innovation and the working of the economy. Finally, despite the aversion of economics to ‘general’ theories, we offer some observations on the support that the user innovation theory of the *numerus clausus* finds in the work of some economic theorists. In particular, the explicit consideration of asymmetries in motivation and cognition, rather than merely information, is in full accord with the strong (yet largely ignored, perhaps even by Coase himself44) admonition of Ronald Coase that economic theories should be judged, not on their ability to predict outcomes, but on the reasonableness of their assumptions.45 Finally, the legal preference for exclusion finds striking resonance in the work of economist Friedrich von Hayek: namely, in his staunch belief that the key to economic success lies in increasing the choice set of the individual.46

44 At least when it comes to the insights about human behavior provided by behavioral economics.
45 Ronald Coase, *The Nature of the Firm: Meaning*, 4 J.L. ECON. & ORG. 19 (1988) at 24: “Most readers will pass over these opening sentences (Putterman omits them when reprinting my article), and others will excuse what they read as a youthful mistake, believing, as so many modern economists do, that we should choose our theories based on the accuracy of their predictions, the realism of the assumptions being utterly irrelevant. I did not believe this in the 1930’s and, as it happens, I still do not. In my Warren Nutter lecture, delivered in 1981, I argue that economists do not, could not, and if they could, should not choose their theories on the basis of the accuracy of their predictions.”
46 See generally F. A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519 (1945). Although the *numerus clausus* restricts the choice set of non-possessors of assets, it increases the choice set of possessors of assets. See Chapter 1, Part 3, 1) *An Alternate Definition: A Legal Preference for Rules of Exclusion*. Because the possessors of assets have advantages in innovating via the use of assets - see generally Chapter 2 – increasing the freedom of possessors at the expense of non-possessors is fully in line with Hayek’s observations about the importance of harnessing the localized knowledge in society by facilitating individual choice.
Chapter 1: The Numerus Clausus -
The Law’s Preference for Exclusion

“It is important to note… that the function of formal rules is to promote certain kinds of exchange but not all exchange.”

---

1 DOUGLASS NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE 47 (Cambridge University Press 1990) (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 law systems. Nonetheless, it is a legal truism that even in the common law, property is significantly less customizable than contract. To determine the metes and bounds of a contractual arrangement, 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

---

3 Merrill & Smith, supra note 2 at 4.
4 Id. at 3-4.
5 Merrill & Smith provide a helpful example of how this procedure of property/contract classification works in practice: “[I]t is useful to consider an example of the principle in operation. Landlord-tenant law includes a version of the *numerus clausus* principle. Leases are limited to four recognized types: the term of years, the periodic tenancy, the tenancy at will, and the tenancy at sufferance. Suppose a landlord and tenant decide to enter into a lease that does not conform to any of the four standard types - a tenancy "for the duration of the war" being the classic example. If landlord-tenant law were just like the law of contract, then there would be no reason not to enforce this agreement in accordance with its terms; that is, the tenancy would last until the war ends. But courts typically do not proceed this way. Instead, they seek to determine which of the four recognized types of leases best fits what the parties have created. Since a term of years requires a "definite calendar ending," and wars last for an uncertain length of time, most courts have concluded that a tenancy "for the duration of the war" must be either a periodic tenancy (if the lease provides for payment of rent at periodic intervals) or a tenancy at will. The result of the pigeon-holing exercise in this example is thus that the parties' intentions are frustrated, because neither a periodic tenancy nor a tenancy at will has the same security of tenure as a tenancy for the duration of the war presumably would have if enforced according to its terms.” (Citations removed.) Id. at 11-12.
servitudes to see whether the legal requirements of ‘touch and concern’, and other such limiting doctrines, are met. If not, then the attempted ‘property’ transaction either fails altogether, or is cast into the rough seas of contract law, where contract’s own set of substantive requirements (most notably consideration) might not be met. And even when such a transaction is enforced as a matter of contract law, the parties’ intention to create a ‘right against the world’ is stymied by the requirement that there be privity between all legally bound parties.

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.’ 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. 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.

It is critical to note that this thesis does not assert that a certain ‘magic’ number of property forms should be allowed – we will never know if the *numerus clausus* is ‘too open’ or ‘too closed.’ Indeed, it is recognized that judicial and legislative recognition of new forms has occurred in the past, and that such innovation continues to occur, particularly in the arena of intangibles. The most notable modern additions to the *numerus clausus* are the right of publicity and the tort of misappropriation of information. Although the right of publicity has been enthusiastically embraced by many state lawmakers and courts, it continues to

---

6 This is essentially equivalent to the perspective adopted by Merrill & Smith, *supra* note 2 at 11: “common-law courts follow the *numerus clausus* as a norm of self-governance.”

7 Tulk v. Moxhay, 41 Eng. Rep. 1143, 1145 (Ch. 1848), marked a major historical milestone in the judicial recognition of new property forms. In this case, the court resorted to equity to circumvent the law’s restrictive attitude towards the recognition of new property forms: “With respect to the observations of Lord Brougham in *Keppell v. Bailey*, he never could have meant to lay down that this Court would not enforce an equity attached to land by the owner, unless under such circumstances as would maintain an action at law. If that be the result of his observations, I can only say that I cannot coincide with it.”

8 As noted by Merrill & Smith, *supra* note 2, at 19 (citations removed): “The *numerus clausus* is probably at its weakest in the area of intellectual property. To be sure, there is considerable stability in the recognized forms of intellectual property, and federal law in the United States preempts many attempts to create novel forms of intellectual property as a matter of state law. But there are some notable exceptions in which judicial creativity in fashioning new intellectual-property interests has been sanctioned.”

9 “[A]bout half the states in recent years have recognized a "right of publicity," which protects the images and voices of celebrities from commercial exploitation without their consent. Slightly more than half of the states
generate debate in both the legal academy and among the judiciary. Misappropriation of information is another example of judicial activism in the arena of intangibles. But this innovation has also been controversial, and unlike the right of publicity, misappropriation of information has been narrowed over time by the judiciary. The controversy that surrounds both of these rights shows how deeply ingrained the *numerus clausus* principle actually is: departures from it generate keen resistance and persistent controversy.

**Part 1 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 recognizing the right have done so at least initially as a matter of judicial lawmaking - which is clearly incompatible with the *numerus clausus* principle.” (Citations removed.) *Id.*

10 See e.g. Eugene Volokh, *Freedom of Speech and the Right of Publicity*, 40 Hous. L. Rev. 903 (2003); and Roberta Rosenthal Kwall, *A Perspective On Human Dignity, The First Amendment, And The Right of Publicity*, 50 B.C. L. Rev 1345 (2009). On occasion, courts have expressed extreme skepticism about the right of publicity. For the most eloquent example of such commentary, see White v. Samsung Elec. Am., Inc., 989 F.2d 1512, 1512-13 (9th Cir. 1993) (citations removed): “Saddam Hussein wants to keep advertisers from using his picture in unflattering contexts. Clint Eastwood doesn't want tabloids to write about him. Rudolf Valentino's heirs want to control his film biography. The Girl Scouts don't want their image soiled by association with certain activities. George Lucas wants to keep Strategic Defense Initiative fans from calling it "Star Wars." Pepsi doesn't want singers to use the word "Pepsi" in their songs. Guy Lombardo wants an exclusive property right to ads that show big bands playing on New Year's Eve. Uri Geller thinks he should be paid for ads showing psychics bending metal through telekinesis. And scads of copyright holders see purple when their creations are made fun of: Something very dangerous is going on here. Private property, including intellectual property, is essential to our way of life. It provides an incentive for investment and innovation; it stimulates the flourishing of our culture; it protects the moral entitlements of people to the fruits of their labors. But reducing too much to private property can be bad medicine.”


12 A comprehensive yet succinct summary of the rise and fall of INS was provided by the court in Barclays Capital, Inc. v. Thellyonthewall.com, 2010 U.S. Dist. LEXIS 25728, 59-66: “After INS, the Second Circuit reacted with hostility to the newly created tort. ... The force of INS as formal precedent was vitiated when most federal common law was abrogated by Erie R.R. Co. v. Tompkins. Nevertheless, INS was adopted into state common law by several states, including most enthusiastically in New York. ... Eventually, the Second Circuit itself reversed course, ruling (over Judge Learned Hand's dissent) that the misappropriation tort was not preempted by the 1909 Copyright Act. New York courts continue to recognize a broad tort of misappropriation to this day as part of unfair competition law. ... The question of whether a hot-news misappropriation claim survives federal preemption was finally resolved in National Basketball Association v. Motorola. ... Under NBA, the elements of an INS claim surviving federal preemption are: (i) a plaintiff generates or gathers information at a cost; (ii) the information is time-sensitive; (iii) a defendant's use of the information constitutes free riding on the plaintiff's efforts; (iv) the defendant is in direct competition with a product or service offered by the plaintiffs; and (v) the ability of other parties to free-ride on the efforts of the plaintiff or others would so reduce the incentive to produce the product or service that its existence or quality would be substantially threatened.” (Citations removed.)
constructed by Merrill and Smith\textsuperscript{13}, and the ‘costs of verifying divided rights’ rationale offered by Hansmann and Kraakman\textsuperscript{14} – each adopt their own version of the principle. In the Merrill and Smith framework, the \textit{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) 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 \textit{sui generis} forms).\textsuperscript{15}

Why is this limited menu of forms necessary? Merrill and 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.”\textsuperscript{16} 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 \textit{numerus clausus}, third parties would potentially have to ‘learn something new’ each time they came across a piece of property. And even if the property in any one case was not a new form, they would have to investigate just to be sure.

\textsuperscript{13} See generally Merrill & Smith, supra note 2.
\textsuperscript{14} See generally Hansmann & Kraakman, supra note 2.
\textsuperscript{15} Merrill & Smith, supra note 2 at 12-20.
\textsuperscript{16} Id. at 35.
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.  

The *numerus clausus* therefore seeks to balance the

---

17 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, *supra* note 2 at 38. For this reason, the theory is often called the ‘communication and third party information costs theory.’ I refer to the theory as the ‘third party information costs theory’ because the term ‘communication’ in Merrill & Smith’s theory sometimes, but not always, refers to communication among the transacting parties. Use of the term can therefore be somewhat misleading.
costs of identifying 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 and Kraakman. Their ‘costs of verifying divided rights’ theory 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 and 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).18 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.”19 In applying this civil law principle to the common law, Hansmann and 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.”20

In Hansmann and 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

---

18 Hansmann & Kraakman, supra note 2 at S374-5: “The degree of notice required and the extent to which the law affirmatively facilitates the giving of notice vary across different types of property rights according to the utility of the partitioning and the costs of giving notice.”
19 Hansmann & Kraakman, supra note 2 at S375. Of course, even the civil law allows some division of the rights in the asset, and Hansmann & Kraakman make note of these: “exceptions included . . . 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.
20 Id. at 376.
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 wish to engage in. If I want to buy a painting that you ‘own’, and the law requires that ownership include the right to resell an 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. According to Hansmann and Kraakman’s theory, therefore, the law 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*.

The ‘third party information costs’ and the ‘costs of verifying divided rights’ theories have much in common. There are, however, several differences between them, and Hansmann and Kraakman carefully catalogue and analyze them. For our purposes, the most

---

21 *Id.* at 375: “Because the benefits of partial property rights are often low and the costs of verifying those rights are generally high, property law necessarily takes an unaccommodating approach to all but a few basic categories of partial property rights.”

22 See Hansmann & Kraakman, *supra* note 2 at S374, describing the theory proposed by Merrill & Smith:

(1) Property rights differ from contract rights by being “good against all the world.” (2) Both the common law and the civil law limit property rights to a small number (*numerus clausus*) of well-defined types. (3) These limits serve to reduce the information-processing costs of potential purchasers of property rights in understanding precisely what they are buying and of potential tortfeasors in understanding with which property rights they must not interfere. (4) [T]here is an “optimal standardization” of property rights, with the appropriate number of alternative forms being determined by a trade-off between the utility of having more forms and the confusion that more forms would engender.

Hansmann & Kraakman differentiate their theory from Merrill & Smith’s, explaining:

(1) Property rights differ from contract rights in that a property right in an asset, unlike a contract right, can be enforced against subsequent transferees of other rights in the asset. That is, a property right “runs with the asset.” (2) The law's limitations on property rights take the form not of standardization into a discreet number of well-defined forms, but rather of regulation of the types and degree of notice required to establish different types of property rights. (3) These limitations serve not to facilitate communication among persons who transact in rights, but rather to facilitate verification of ownership of the rights offered for conveyance. (4) Property law generally addresses the verification
relevant difference between them is that in Merrill and Smith’s view, there is an asserted upper bound on the socially optimal number of property forms, because notice does not completely solve the problem that the numerus clausus is designed to address. 23 In Hansmann and 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. 24

Both of these theories represent significant contributions to the legal corpus. They are not only 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. We will discuss those features in Part 2.

Part 2 WHY WE NEED A NEW THEORY

Although these theories represent unique contributions to the literature, they do have much in common; and for the purposes of this thesis, 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; 25 and 2) notice plays a pivotal role

---

23 Merrill & Smith cite this upper bound as being produced by the intractable features of human ‘communication’. Merrill & Smith, supra note 2 at 45. 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.” Id. at 44. 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.

24 Hansmann & Kraakman, supra note 2 at S416-17.

25 “We agree, in particular, that third-party information costs are central to the law’s regulation of property rights.” Hansmann & Kraakman, supra note 2 at S374. 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
in both theories in determining the need for the *numerus clausus*\(^\text{26}\). 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 law is actually concerned with what goes on between the parties to the transaction, and because it resists making the *numerus clausus* turn on notice alone. We will discuss each of these in turn.

1) The ‘Third Party’ Problem

The first objection to these two theories of the *numerus clausus* is that they focus on the costs that new property forms impose on third parties, rather than on persons within the chain of asset transfer. The focus on third party costs represents an uncomfortable fit with the structure and goals of private law, which generally seeks to adjudicate the rights as between the individuals before the court\(^\text{27}\). Although there is certainly a multifaceted policy background that informs and shapes all of the law, whether public or private, it would seem to be an odd thing for the defining feature of property law – the *numerus clausus* - to have *nothing* to do with the actual owners and possessors of that property.

In the case of Merrill and Smith’s theory, there are additional objections to its focus on third parties, which arise from that theory’s unique articulation of the third party information costs issue. Recall that Merrill and Smith liken the creation of new property forms to the addition of words to a language. Creating new property forms creates ‘noise’ and confusion for third parties because everyone must now learn about and search through the expanded universe of property rights. However, the ‘noise’ problem only occurs in a

---

\(^{26}\) If third parties can be cheaply, quickly and accurately notified of unusual forms of property, notice can substitute for standardization. Despite Merrill & Smith’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.

\(^{27}\) “The most striking feature of private law is that it directly connects two particular parties through the phenomenon of liability. Both procedure and doctrine express this connection. ... Private law looks neither to the litigants individually, nor to the interests of the community as a whole, but to a bipolar relationship of liability.” ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW 1-2 (Harvard University Press 1995).
select number of cases, because the ‘noise’ that Merrill and Smith hypothesize is created only if individuals actually see or encounter the other, newly created property types. In most cases a third party will never even be aware of other property types that exist; he will simply deal with the property over which he is seeking to transact. And even if there were only a few property types available, an individual would still have to check to see what type of property he is actually dealing with - so that is a constant cost. It is not clear, for example, why encountering an unusual easement in a deed is any more of an informational burden than encountering an easement that was selected from a limited list of options. Finally, 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, because these are precisely the cases in which property is highly customizable.

2) The ‘Notice’ Problem

Notice plays an important role in both Merrill and Smith’s and Hansmann and Kraakman’s theory of the *numerus clausus*. Therefore, in both theories, the availability of land registries and other methods of recordation does much to determine the boundaries of the principle. However, as we saw in Part 1, notice is far more pivotal in the ‘costs of verifying divided rights’ theory. Thus, this thesis’ treatment of notice diverges more sharply from that in Hansmann and Kraakman’s theory than in Merrill and Smith’s theory. For this reason, we focus our comments about notice on Hansmann and Kraakman’s theory, even though the general thrust of these criticisms applies equally to both.

Recall that the ‘costs of verifying divided rights’ 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

---

28 This objection is shared by Hansmann & Kraakman, supra note 2 at S381 and S401.
29 Robinson makes a similar point. See Glen Robinson, *Personal Property Servitudes*, 71 U. CHI. L. REV. 1449 (2004) at 1486-7: “What they overlook is the fact that if every buyer must be given specific notice of any deviation from the baseline of full title transfer, the information cost problem is solved; there is no externality. A legally required notice has the same effect of internalizing information costs as a legal guarantee of formal title under a *numerus clausus* principle.”
30 See e.g. Merrill & Smith, supra note 2 at 40: “As the costs of standardization... shift, we expect the optimal degree of standardization to rise or fall. Consider the rise of registers of interests in real property, that is, recording acts. This device lowers the costs of notice.”
31 *Supra* notes 23-24 and accompanying text.
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,\(^{32}\) in which the court decided that the principle of patent exhaustion allowed a buyer to use a lawfully manufactured device without the further permission of the patent owner.\(^{33}\) That is, once the patent was lawfully embodied in a device and that device was sold, the rights in that device (i.e., that particular embodiment of the patent) were allocated to the buyer (i.e., the rights in the *device itself* were *not* divided between the patent holder and the buyer). As a result of the court’s refusal to divide rights in this case, a licensed manufacturer of chipsets and microprocessors (Intel) could sell the devices to computer manufacturers, and the computer manufacturers (Quanta) did not need to pay the patent holder (LG) for a license to use the device in assembling its computers.

It should be noted that both the cost of providing notice, and the value of the prospective divided right, were irrelevant to the outcome of this case. There was no question in the LG case that all parties knew of the patent holder’s assertion that purchasers of the lawfully manufactured devices were not licensed to use the devices in assembling computers containing non-Intel components - so there was no notice problem. In addition, the value of the case was substantial: upwards of several million dollars for LG alone,\(^{34}\) not to mention the other patentees who would have benefited from the ability to claim royalties from downstream buyers of patented devices. Thus, the *Quanta v. LG* case shows that, contrary to Hansmann and Kraakman’s theory, the law *prevents* the creation of divided rights even when notice is cheap and effective, and when the value of the prospective divided rights is high. Surprisingly, Hansmann and Kraakman themselves admit that courts refuse to enforce divided rights even where there is no notice problem, particularly in the personal property cases: “the restrictive covenants in these cases imposed few problems of notice, and their rejection on that ground would be hard to justify.”\(^{35}\) However, like Robinson, they attribute

---


\(^{33}\) Although of course it needed the permission of the patent owner to make the device in the first place.

\(^{34}\) It is impossible to determine precisely what the divided right in this case would have been worth, because this requires that we estimate what damages and future royalties a court might have awarded to LG. This task has been likened to an act of ‘crystal ball gazing’ because is difficult to determine how much value the patents actually contributed to Quanta’s end product the (the assembled computers). See Richard Stern, *Quanta Computer Inc v LGE Electronics Inc – Comments on the Reaffirmance of the Exhaustion Doctrine in the United States*, 12 E.I P.R. 527 (2008) at 534.

\(^{35}\) Hansmann & Kraakman, *supra* note 2 at S417.
the outcome in the personal property cases (which are at the core of the *numerus clausus* puzzle) to reasons of public policy which are unrelated to their theory (because they have nothing to do with the cost of notice or the verification of rights).  

While it is true that a court’s refusal to enforce divided rights can sometimes be understood in terms of some independent public policy (the price maintenance cases fall into this category, because they can be understood as being in restraint of trade), in other cases it is not apparent what such policies might be. In fact, the absence of such sound public policy reasons is what gives the criticisms of scholars like Robinson so much persuasive force. In other words, if a court’s refusal to enforce a divided right is a product of some loose grab bag of public policies which are themselves unrelated to property law, then there is no justification for such refusal if we cannot identify what that policy is. Even more troubling, a public policy version of the *numerus clausus* is far narrower than the one that currently exists in the decisions, because most of the plausible public policy goals against divided rights boil down to one: restraint of trade. And as Robinson so rightly notes, the weight of theoretical and empirical evidence is beginning to turn against even those practices that were once considered to be within the core of antitrust theory, such as resale restrictions and price maintenance schemes.

Hansmann and Kraakman have an interesting way of dealing with the apparent challenge that the personal property cases pose for their theory. They simply assert that there is no *numerus clausus*, as a matter of property law, at all: “it presumably remains open to a court to decide [to enforce divided rights in personal property]… .There is no particular reason to have a true *numerus clausus* in the sense of a formal doctrine that explicitly limits property rights to a small well-defined set and requires that all other rights be based on

---

36 “[I]t is understandable that the common law has no formal doctrine equivalent to the civil law’s *numerus clausus* and accompanying unitary theory of property rights. The occasional judicial pronouncements to the contrary frequently have the appearance of makeweight efforts to bolster holdings that are at bottom motivated by narrower considerations of policy. For example, the supposed doctrine that U.S. law will not permit creation of a servitude on a chattel has been invoked most conspicuously in cases involving manufacturers’ efforts to impose restrictions on the use of their products, such as resale price maintenance. The courts’ rejection of these efforts was evidently motivated by considerations of competition policy that had nothing to do with the problems of notice that generally lie behind the law’s unaccommodating approach to divided property rights.” *Id.*

37 *See e.g.* Dr. Miles Medical Co v John D. Park & Sons Co, 220 US 373 (1911).

38 Robinson, *supra* note 29 at 1502-1503.
explicit contracts.” Thus, the most interesting court decisions – those that deal with personal property and intellectual property - are relegated to that murky theoretical wasteland known as ‘public policy.’ Indeed, after having spent most of their paper supposedly constructing a theory of the *numerus clausus*, at the end of their paper Hansmann and Kraakman go so far as to assert that, in reality, there is no *numerus clausus* in the civil law either. Instead, they locate the principle in the politics of the French Revolution, and the subsequent development of legal scholarship and theory.

Although it may seem surprising that a theory would deny the existence of the very thing that it purports to explain, Hansmann and Kraakman’s assertion that there is no *numerus clausus* as a matter of property law (rather than as a creature of public policy or political history) is actually consistent with their theory. That is to say, Hansmann and Kraakman are absolutely correct that a *numerus clausus* that turns solely on notice is equivalent to no *numerus clausus* at all. From the point of view of the user innovation theory of the *numerus clausus*, that is precisely the problem.

Part 3  
**CONSTRUCTING A NEW THEORY**

In this Part, we begin our construction of a new theory of the *numerus clausus* - one that accounts for the law’s focus on the parties to the transaction, and for its refusal to let notice determine the boundaries of the doctrine. First, we will formulate a new definition of what the *numerus clausus* is. Then, we will briefly discuss the major question which any theory of the *numerus clausus* must answer: why don’t the property rights of prior owners trump the property rights of downstream owners?

---

40 Id. at S417.
41 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 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.
1) An Alternate Definition: A Legal Preference for Rules of Exclusion

In this thesis, 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). As used in this thesis, the term ‘exclusion’ largely coincides with the meaning that ‘private property’ has in both the legal and economic literature; whereas the use of the term ‘governance’ diverges from its current usage. 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,’ 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 it a full set of legal privileges (use), powers (transfer) and immunities (income). Under a rule of exclusion, in other words, full and free use travels with permanent possession of an asset. This means that, as a property owner, you neither have to ask permission for, nor make payment for, the use of your property.

It is on the governance side of the terminology that this thesis adopts a specialized meaning. Under a governance rule, the possessors of assets must either obtain permission to use an asset, or they must pay for such use (or both). Under a governance rule, therefore, rights of use in an asset are at least partially separated from the right to possess that asset. This definition of governance largely coincides with Hansmann and Kraakman’s use of the

---

42 See e.g. E.C.S. Wade, *Restrictions on User*, 44 L. Q. REV. 51 (1928).
44 ‘Permanent’ is used as a qualifier here to distinguish changes in possession that are merely temporary (e.g., rental of a chattel).
45 ‘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. See Chapter 4.
46 It should be noted, however, that Hansmann & Kraakman’s use of the term ‘divided rights’ differs somewhat from the use of the term in this thesis. Throughout their paper, they refer to divisions of use and possession, but
term ‘divided rights;’ therefore, the terms ‘governance’, ‘divided rights’, ‘use restrictions’, and ‘separations of use and possession’ will be used interchangeably throughout this thesis. This thesis does not simply use the term ‘divided rights’ because the term ‘governance’ can be understood to coincide with the term ‘divided rights’ in all circumstances except one: self-enforcing technologies. When governance rules are implemented via self-enforcing technologies (such as computer code), it makes no sense to speak of ‘divided rights’, ‘use restrictions’ or ‘separations of use and possession’ because the governance rule becomes part of the product itself. That is, the governance rule becomes an inherent and inseparable characteristic of the asset itself.

The use of the term ‘governance’ in this thesis differs from its usage in the legal and economic literature. There, the term governance refers to those situations in which: 1) the uses of property are specified, or 2) in which there is shared access to property. In contrast, in this thesis the term governance refers to any division of control or income across different individuals. Governance applies to private ownership just as much as it applies to common ownership schemes; and it neither requires shared access, nor the specification of uses. Servitudes, which separate the possession of an asset from certain use rights in that asset, are the canonical example of governance rules, because they restrict the ability of an owner to unilaterally and freely determine the uses of his property. Thus, the law’s antipathy to servitudes is an aversion to governance rules.

Conceptualizing the numeros clausus as the law’s means to mediate between governance and exclusion allows us to isolate the relevant characteristics of a range of seemingly unrelated phenomena: the enforcement of standard form contracts, the role of notice in drawing the line between property and contract, the creation of novel servitudes and near the end of their paper they equate ‘divided rights’ with ‘fragmented rights.’ See Hansmann & Kraakman, supra note 2 at S415. In this thesis, the distinction between divided rights and fragmented rights is critical. See Chapter 3, Part 3, Divided Rights Versus Fragmented Rights.

48 See e.g. Henry E. Smith, Exclusion Versus Governance: Two Strategies for Delineating Property Rights, 31 J. LEGAL STUDIES S453 (2002) at S455: “At the pole opposite exclusion along the organizational dimension are what I am calling governance rules. These rules, more identified with common property in the literature, pick out uses and users in more detail.”

49 See e.g. C. Leigh Anderson & Eugene Swimmer, Some Empirical Evidence on Property Rights of First Peoples, 33 J. ECON. BEHAVIOR & ORG. 1 (1997) at 10: “Governance costs only arise under the shared access regimes.”

50 Actually, it is more accurate to say that as used here, the term governance refers to any division of control or income, or any separation of residual control or residual income. See infra Chapter 2, Part 1, Departures From Previous Theories.
other forms of property, and the technological implementation of use restrictions on personal property (and the legal enforcement of these technological measures). These phenomena are closely related because they are all about one thing: governance. But governance is a concept with many faces - some of which may appear to be completely unrelated - because governance has many methods of implementation. Governance rules can be created by contract (you agree to use restrictions); by notice (the law requires observance of use restrictions because you knew of them and it would be unfair of you to ignore them); by property (the use right was not transferred to you); and by technology (the use is restricted by technological controls, product design, or inherent characteristics of the tangible). The realization that governance comes in many guises is the key to understanding the link between the broad range of phenomena - from controversies about digital rights management in copyright law, to concerns about mass market contracting, to the resistance to servitudes on chattels - that we analyze in more detail in Chapters 4 and 5.

The governance/exclusion terminology may be unfamiliar to some legal readers, and may therefore lack the intuitive familiarity of other terminology. However, these terms are nonetheless superior to all of their near substitutes, of which there are four: contract/property, in personam/in rem, license/sale, and lease/sale. The first such substitute, ‘contract or property,’ is the least satisfactory. The reason for this is simple: the term ‘property’ can refer to a servitude which travels with an asset, and as we have already seen, a servitude is the paradigmatic governance rule. So, although contracts are governance rules, the term property does not always refer to rules of exclusion.

A closely related distinction, in personam or in rem, has elsewhere been equated to the governance/exclusion distinction (i.e., it has been stated that in personam rights represent governance rules and in rem rights represent exclusion rules). Although this is true in the majority of cases, it is not always true. A contract is the paradigmatic in personam governance rule. But servitudes that run with an asset are in rem governance rules. Thus, the in personam/in rem language does not correspond to the governance/exclusion

51 See Thomas Merrill & Henry Smith, The Property/Contract Interface, 101 COLUM. L. REV. 773 (2001) at 791: “Viewed in this way, in personam contract rights and in rem property rights are species of two more general strategies for regulating resource use: In personam rights are an instance of what can be called a governance strategy for determining use rights; in rem rights reflect an exclusion strategy for determining use rights.”

52 Servitudes that require notice are nominally in personam, but they function like in rem rights if notice literally ‘travels with the asset.’
termology. Indeed, the problem which the numerus clausus seeks to address is precisely the overlap between governance rules and in rem rights. It would therefore be counterproductive to adopt a terminology which obscures the very problem that we are trying to address. In addition to these problems, there is a related disadvantage to the in personam/in rem distinction: it loses its coherence in the face of notice. Notice that travels with an asset bridges the gap between in personam and in rem rights, transforming in personam obligations into de facto in rem obligations.53

A somewhat closer alternative to the governance/exclusion terminology is the ‘license or sale’ distinction. Although this distinction has the advantage of being familiar to many readers, it suffers from an unacceptable degree of ambiguity, because the term ‘license’ can refer to either a contractual transaction or a property transaction. Although licenses are generally interpreted as contractual in American law,54 it is frequently asserted that a license is not a creature of contract, but is rather a ‘unilateral abrogation of rights.’55 If a license is a creature of contract, then the law has a robust set of tools (agreement, consideration, etc.) with which to police the availability of governance rules. If a license is a unilateral abrogation of rights, on the other hand, then it can be used as an open-ended justification for the implementation of governance rules.

Perhaps the closest alternative to the governance/exclusion terminology is the choice between a lease and a sale. A lease “involves possession or use for a limited time;”56 whereas a sale involves permanent use and possession, and a transfer of title.57 Indeed, it is the permanence of the possession which determines the transfer of title. The lease/sale terminology is the closest alternative to the governance/exclusion terminology because the phenomenon that the numerus clausus represents can be roughly reduced to a single question: Why doesn’t the law recognize permanent leases?58 Why, in other words, does a

53 This phenomenon reaches its height when such notice is technologically implemented, and terms are incorporated into the good itself. See e.g., Ariel Katz, Substitution and Schumpeterian Effects Over the Life Cycle of Copyrighted Works, 49 JURIMETRICS 113, 143-144 (2009).
54 See e.g. Everex Sys. v. Cadtrak Corp. (In re CFLC, Inc.), 89 F.3d 673 (9th Cir. 1996), analyzing a patent license as an executory contract.
57 Id.
58 Thanks are due to Jeremy Paul for this insight.
lease have to have a term – i.e., why does a leased item have to be eventually returned to the lessor?

Although the lease terminology is more frequently used in the real and personal property context, it is just as applicable to transactions involving transfers of copies of intellectual property. Indeed, one of the best examples of leasing can be found in the movie industry: studios lease copies of movies to theatres, and require their return after their run in the theatre ends. The lease terminology also has the benefit of clarifying the legal status of transactions that permanently transfer copies of intellectual property works. In these transactions, the mantra that ‘this copy is licensed not sold’ is used precisely because it confuses the licensing of the intellectual property work (of course the intellectual property content of the copy can be licensed), with the permanent transfer of the copy (the copy is not itself ‘licensed’ unless it actually needs to be returned – in which case the copy is leased, not licensed). However, although the term ‘lease’ is roughly coincident with an ability to impose a governance regime on a transferred asset, the ‘sale’ side of the equation suffers from the drawback that a servitude can sometimes be imposed on an asset that is ‘sold’ to a buyer. Thus, the term ‘sale’ is not always equivalent to a rule of exclusion.

2) Why Is Prior In Time Not Always Prior In Right?

Now that we have established that the numeros 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? 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 (though less exact) terms. Why does the law

---

60 Id. at 167-169.
61 Typical leases, in which use and possession are transferred in return for periodic payments and the eventual return of the asset, may not sound much like governance rules, since they usually give the possessor full use of the asset. However, the term ‘governance’, as used here, encompasses those situations in which payments but not permissions flow between the possessor and the owner. In other words, governance rules include not only separations of use and possession (where both permissions and payments are required for the possessor to make use of the asset), but also mere open-ended liability rules (where payments, but not permissions, are required in order for the possessor to make use of the asset).
overwhelmingly disfavor use restrictions when a tangible asset 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. 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. 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, and has given rise to a great deal of scholarly criticism. The question

62 Replacing sale with a series of bailment-like arrangements might sound improbable, until we consider that the software industry attempted this very thing, see Thomas M.S. Hemnes, Restraints on Alienation, Equitable Servitudes, and the Feudal Nature of Computer Software Licensing, 71 DENV. U.L. REV. 577 (1994). In their paper, Hansmann & Kraakman discuss two extant examples of divided rights: 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.

63 A close reading of Kepell v. Bailey, 39 Eng. Rep. 1042 (Ch. 1834), 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 about ‘first sale’ in the intellectual property context, Bobbs-Merrill Co. v. Straus, 210 U.S. 339 (1908), merely stated, without explanation, that the imposition of a servitude-like restriction on a book was outside the Congressional copyright grant. This point of weakness in the decision has been noted by other scholars, see e.g. Robinson, supra note 29 at 1470: “The Court in Bobbs Merrill did not articulate any policy basis for its new rule [first sale] beyond simply declaring that a copy-right owner's exclusive right to sell did not authorize restrictions on the resale of the copyrighted work.” Hence, reading the tea leaves of the common law to divine a solid rationale for the law’s preference for exclusion is a largely fruitless enterprise.

64 See e.g. Percy Bordwell, Alienability and Perpetuities II, 23 IOWA L REV 1, 14 (1937): “The great objection to repugnancy as an explanation, however, is that it explains nothing. It begs the question. By definition it makes alienability a characteristic of the fee simple and then rejects general restraints as inconsistent with the definition. Logically any restraint would be as inconsistent as a general restraint and this has often been pointed out. The allowance of partial restraints, therefore, destroys the logic of repugnancy and logic is its one support." Also see Robinson, id. at 1480-81: “But with the feudal system out of the way one might expect some explanation for refusing to allow ordinary property owners to dispose of their property on such terms as they choose. As often as not the "reason" given is simply that it cannot be done because the right of alienability is inherent in the property right being transferred. Coke's formulation of the rule more than three centuries ago explained that when an owner of property sells or gives "his whole interest" in the property he cannot at the same time prohibit its further alienation "because his whole interest and propertie is out of him, so as he hath no possibilitie of a reverter, and it is against trade and traffique and bargaining and contracting between man and
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. 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 the work of Richard Epstein. 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 should trump the first owner's rights to burden the property.” This is the challenge that any theory of the numerus clausus must answer.

The answer to this legal puzzle is found in the deleterious impact that divided rights have on user innovation. We will discuss the phenomenon of user innovation in detail in
Chapter 2. There, we will see that the incidence of user innovation is lower when rights of use are separated from the right of possession because property rights allocations themselves change human behavior. Individuals both \textit{interact} differently with assets, and \textit{transact} differently over them, depending on how rights of use and possession in those assets are divided. The changes in human behavior that accompany property rights allocations mean that both \textit{resource values} change and \textit{transaction costs} change as property rights allocations change. Resource values are higher when rights of use and possession are unified because a significant amount of valuable innovation can be revealed only through asset use. Possessors interact differently with assets than non-possessors, and users have advantages over nonusers in seeking out new uses of assets that can be revealed only through asset use. This differential interaction with assets is critical to entrepreneurship and wealth creation. Furthermore, because each individual has a unique stock of rival and non-transferable human capital, each possessor interacts differently with an asset than every other possessor of a duplicate asset. This means that the results of user innovation cannot be contracted for in advance (\textit{ex ante}).

Transaction costs also vary with property rights allocations. Most importantly for us, divided rights dramatically decrease the willingness to deal of \textit{both} the owner of the use right, and the possessor (user) of the asset.\textsuperscript{68} As a result, what at first appear to be situations of low transaction costs (e.g., a buyer has permanent possession of an asset, a seller retains use rights, and technology allows low-cost monitoring, enforcement and payment for incremental use), become situations of high transaction costs \textit{by virtue of the property rights allocations themselves} (e.g., the buyer’s possession of the asset gives rise to subjective feelings of ownership, which results in attempts to disable the technological mechanisms). Thus, \textit{ex post} rearrangement of divided rights will often not occur, even when a mutually beneficial transaction would otherwise result. The difficulties of contracting around governance rules will be discussed in detail in Chapter 3.

\footnotesize{\textsuperscript{68} This occurs because \textit{both} possession of an asset and legal ownership of use rights in an asset give rise to feeling of psychological ownership, and individuals will resist having to bargain for something that they feel they already ‘own.’ \textit{See generally}, Chapter 3.}
In this Chapter, we examined the two major economic theories of the *numerus clausus*, and found that they share two characteristics which do not coincide with the treatment of the principle by the courts. First, these theories focus on the impact that novel property forms have on third parties, whereas the law focuses on the parties to the transaction. Second, in these theories the *numerus clausus* opens up with notice, whereas the courts refuse to let notice determine the boundaries of the doctrine.

We began our search for a theory which would account for these features of the law by offering a new definition of the *numerus clausus*. This definition equates the *numerus clausus* with a general legal preference for rules of exclusion (and a concomitant aversion to governance rules). At first glance, the legal preference for exclusion seems inexplicable, because it flies in the face of the principle that the rights of prior property owners should trump the legal rights of downstream possessors. We asserted that the justification for this peculiarity is two-fold: rules of exclusion foster user innovation (which raises resource values); and individuals will not contract around governance rules (because feelings of ownership arise from both use/possession *and* legal ownership). Exploring these assertions in more detail will be the respective tasks of Chapters 2 and 3.
Chapter 2: User Innovation – Why Exclusion is Important

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

---

1 ERIC VON HIPPEL, DEMOCRATIZING INNOVATION 2 (MIT Press 2005).
INTRODUCTION

In this Chapter, we will begin our consideration of why the law should prefer exclusion over governance. The answer to this question has already been briefly alluded to in Chapter 1: exclusion should be preferred over governance whenever it allows for a higher-value use of a burdened asset.\(^2\) That is to say, exclusion will be superior to governance whenever the possessors of assets seek out new, higher-value uses of those assets. As we shall see, possessors of assets and owners of use rights do not have the same ability to search out new uses, because many new uses can be revealed only through actual use of an asset. People will interact differently with assets depending on how rights of use and possession are divided, and this means that exclusion and governance do not produce equivalent innovative outcomes.

We will also see that the innovative impact of unifying use and possession increases with the number of duplicate assets released into the market, because each possessor of an asset represents a unique set of innovative opportunities. In other words, each possessor of an asset interacts differently with an asset than every other possessor, because each possessor has a unique stock of human capital. This means that the impact of exclusion will be particularly important in the case of non-unique chattels, and intellectual property which is sold in multiple copies. We will refer to the phenomenon whereby the possessors of assets reveal new uses, or produce new innovations, during the course of asset use as user innovation.

Part 1 DEPARTURES FROM PREVIOUS THEORIES

The user innovation theory of the numerus clausus differs from traditional economic analyses in three main ways: 1) it deals with first party rather than third party effects; 2) it deals with the immediate present rather the far future; and 3) it deals with results that cannot be contracted for \textit{ex ante}. We will briefly discuss each of these differences in turn.

\(^2\) In this chapter, we will discuss the \textit{ex ante} innovative benefits of exclusion – i.e., how exclusion fosters user innovation by promoting the innovation that occurs as a by-product of asset use. This type of innovation cannot be contracted for in advance. Of course, if individuals contracted around governance rules \textit{ex post} to the innovative act, governance rules would duplicate the advantages of exclusion. We will discuss why parties will frequently not contract around governance rules in Chapter 3.
i) First Party Versus Third Party Effects

The focus on user innovation gives rise to the first major departure from previous theories of the *numerus clausus*, as it brings us squarely within the chain of privity of the asset (i.e., those to whom the asset is transferred, either now or in the future). Third party concerns play no role in the user innovation story, and so they play no role in the user innovation theory of the *numerus clausus*.\(^3\) This is in contrast to the two extant theories, which focus on the information costs to third parties, rather than to parties within the chain of title or privity.\(^4\) 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 – those who are not directly involved with the asset in question.\(^5\) The reason for this is straightforward: the costs to persons within the ‘chain of privity’ are generally considered to be accounted for by the price mechanism.\(^6\) That is to say, the more restrictions there are on a piece of property (i.e., the more robust the

\(^3\) Although user innovation has a greater impact on resource values precisely because it creates spillovers, the *numerus clausus* contributes to the downstream creation of spillovers precisely because it focuses on the parties to the transaction. In other words, user innovation is (in part, but not wholly) about spillovers, but the user innovation theory of the *numerus clausus* is about the bottleneck to those spillovers. All the problems discussed in Chapter 3 – psychological ownership, etc. – are also first party issues. In fact, the reason why previous theories of the *numerus clausus* have been unable to adequately account for the *numerus clausus* is because economic theory moves too quickly from first party to third party issues. When we consider the fact that technological developments blur the line between first party and third party issues – i.e., that notice that travels with an asset transforms all parties who come into contact with the asset into a first party – it becomes even more important to focus on first parties. Modern scholars are particularly attuned to the slippery nature of the line between first and third parties, see e.g., Ariel Katz, *Substitution and Schumpeterian Effects Over the Life Cycle of Copyrighted Works*, 49 JURIMETRICS 113, 143-144 (2009); and Margaret Radin, *Humans, Computers, and Binding Commitment*, 75 IND. L.J. 1125 (2000) [Radin, Commitment]. But this awareness has been part of the common law since at least 1834, see Keppell v. Bailey 39 Eng. Rep. 1042, 1053 (Ch. 1834): “Every assignee of a lease has notice of the lessor's covenants; consequently no covenant, however absurd soever, 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.”

\(^4\) See e.g. Henry Hansmann and Reinier Kraakman, *Property, Contract, and Verification: The “Numerus Clausus” Problem and the Divisibility of Rights*, 31 J. Legal Studies S373, S374 (2002) [Hansmann & Kraakman]: “We agree, in particular, that third-party information costs are central to the law’s regulation of property rights.”

\(^5\) Id.

\(^6\) See e.g. Thomas W. Merrill and Henry E. Smith, *Optimal Standardization in the Law of Property: The Numerus Clausus Principle*, 110 YALE L.J. 1, 30 (2000) [Merrill & Smith]: “the costs to potential successors in interest will also be mediated through the price mechanism and so will not require legal intervention.”
governance rule is), the lower its price will be. Thus, the price of the asset will reflect all the costs that the idiosyncratic property form imposes on all those within the chain of privity.

According to the canonical view, therefore, any rationale for the *numerus clausus* must look *outside* the chain of privity for its justification (i.e., 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. In short, when property law replaces a rule of governance with a rule of exclusion, it *supersedes the price mechanism.*\(^7\) Consider a governance rule on a transferred asset which allows a specified use of the asset, but which requires a series of incremental permissions and payments for any additional uses which may arise. Such a scenario would occur, for example, with a chattel that is sold to an end user, but which is affixed with an enforceable ‘use on Tuesdays only’ label. If the user wants to use the asset on Mondays and Wednesdays as well, he must return to the seller and negotiate the terms of such use. However, if the law disallows this governance regime and replaces it with a rule of exclusion, this series of payments and permissions (contracts for incremental use) is replaced by property ownership (a single contract for sale).\(^8\) In addition, the imposition of a rule of exclusion means that some of the future uses made by the asset possessor (now owner) are unpriced (they are not directed by the price mechanism, but by the will of the property owner).

Of course, confusion about the relation between price and property inevitably arises because the expected uses at the time of the sale often affect the price at which the property is ‘sold’.\(^9\) If a seller suspects that a buyer will get greater use of the property, then he is more likely to demand a higher price (though his ability to actually receive this price will be affected by the availability of competitive alternatives). Similarly, the price that the buyer is willing to pay can be expected to increase as the value of his expected uses increase. But

\(^7\) The relevance of the supersession of the price mechanism was first explored in terms of the firm, *see generally* Ronald Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937) [Coase].

\(^8\) Compare this to the firm, where a series of product markets between the input owners and the end consumer is replaced by a series of factor markets between the input owners (who are now employees) and an agent (the firm). Similarly, the complex series of contracts between each of the factors of production that would be necessary to coordinate production without the firm, is replaced by a much lower number of contracts of employment between each input owner and the firm. For an insightful discussion of the difficulty in defining actually what it means for the firm to supersede the market *see* Steven N. S. Cheung, *The Contractual Nature of the Firm*, 26 J.L. & ECON. 1, 3-9 (1983).

\(^9\) Similarly, when a firm replaces a series of products markets with a series of factor markets, the firm inputs (the employees, etc) are priced (remunerated). In other words, the firm supersedes the price mechanism, but it is still “related to an outside network of relative prices and costs” Coase, *supra* note 7 at 389 (quoting L. Robbins).
when considering whether the transfer of an asset supersedes the price mechanism, the
question is not whether the price at the time of transfer (t1) accounts for the full expected
value that the buyer will extract from the asset (though it is reasonable to assume that there
must be some difference between this and the actual price for a transfer to occur). Instead, the
question is whether an unexpected change of value \(^{10}\) some time after the transfer (at t2) will
accrue to the buyer or to the seller. 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 the price mechanism 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 some 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.\(^{11}\) More accurately, the term ‘exclusion,’ as used
in this thesis, is equivalent to ownership of residual control \textit{and} residual income.\(^{12}\)

Another way to see that a rule of exclusion supersedes the price mechanism is to
compare property ownership to a firm. In a firm, resource use is directed by the controlling
mind of the firm owner, rather than by market arrangements. Similarly, under property
ownership, resource use is directed by the controlling mind of the property owner, rather than
by contractual arrangements. If I own property, I do not have to pay others for the right to use
it, and I do not have to ask others for their permission to use it. I am no longer subject to any
governance rules that a former owner might wish to impose on my use of the asset. The \textit{numerus clausus}, because it replaces a governance rule with a rule of exclusion, thus operates
to supersede the price mechanism: that is its purpose. Supersession of the price mechanism -
the replacement of governance by exclusion - is the very essence of legal ownership. If we
accept the assertion that the hallmark of property is the supersession of the price mechanism,

\(^{10}\) This change in value could easily be negative, but for ease of discussion we will assume an increase in value. Sellers would obviously have no interest in trying to ‘capture’ (i.e., pay for) unexpected decreases in value post-transfer. Although sellers and buyers try to anticipate future changes in value when agreeing on a price, the assumption that a price actually factors in all such future changes is little more than a convenient fiction. The principle that ‘nobody knows’ - which has been grudgingly accepted by most people in the creative industries, see Richard E. Caves, Creative Industries 3 (2000) – actually applies to everything in life. Nobody knows what will happen, what will work and what will not, what the future value of a resource will be, who will innovate and who will not, etc..

\(^{11}\) See e.g. Oliver Hart, Firms, Contracts and Financial Structure 30 (Oxford University Press 1995) [Hart].

\(^{12}\) For a discussion of residual control and residual income see id. at 63-66.
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? In other words, why does the *numerus clausus* supersede the price mechanism? The answer to this question, it shall be asserted, is that superseding the price mechanism fosters user innovation.

**ii) The Immediate Present Versus the Far Future**

This theory’s focus on user innovation also marks an important 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.\(^\text{13}\) 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.\(^\text{14}\) 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 restricting 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.


\(^\text{14}\) *Id.*
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. But it is not at all clear that property owners either have the will or the way to perform this calculation, even if private ownership is indeed better than communal ownership in this regard. But focusing on user innovation avoids 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.

### iii) The Impossibility of *Ex Ante* Contracting

Throughout this chapter, we will see that the emergent nature of innovation means that no one (including the innovators themselves) can predict who will innovate, when they will innovate, or how they will innovate. This means that no one (including the innovators themselves) can accurately judge *ex ante* when they should and should not give up their rights to innovate in exchange for a cheaper asset price, or when they should agree to ‘grant back’ such innovations to the seller of the asset. In most cases, it would seem that a rational buyer would opt for a lower up-front price. If such buyers later become innovators, when it

---

15 *See e.g.* Harold Demsetz, *Toward a Theory of Property Rights*, 57 AMER. ECON. REV. 347, 356 (1967) [Demsetz].
17 They may simply not be able to predict future contingencies. *Id.*
18 “Communal property means that future generations must speak for themselves. No one has yet estimated the costs of carrying on such a conversation.” Demsetz, *supra* note 15 at 355.
19 Under a ‘grant-back’ agreement, the buyer agrees to give the seller the right to use improvements made by the buyer to the seller’s product. *See e.g.*, ABA SECTION OF ANTITRUST LAW, *HANDBOOK ON THE ANTITRUST ASPECTS OF STANDARDS SETTING* 74 (2004) [Antitrust Handbook]. They are normally employed when sellers license assets containing intellectual property to other firms, but there is no theoretical reason why they could not also be used when user innovations (whether patentable or unpatentable, copyrightable or uncopyrightable) arise from consumer use of products (whether patentable or unpatentable, copyrightable or uncopyrightable).
comes time for them to reveal their breach of the ‘non-innovation’ agreement, or to ‘grant
back’ their innovations to the seller, the phenomena discussed in Chapter 3 will interfere with
such ex post rearrangements of rights.20

On the other side of the equation, rational sellers would want to have first refusals,
grant-backs, or royalty streams for user innovations that are complementary to their products,
and prohibitions (or grant-backs21) on user innovations that are substitutes for their products.
But how are sellers going to know beforehand which users will produce complements, and
which will produce substitutes? A second major problem involves disclosure: how are sellers
going to provide incentive for users to both disclose the innovation to the seller and refrain
from disclosing the innovation to the public? As we will see below, a significant amount of
user innovation occurs in situations where the use (and even the development) of the
innovation will be naturally observed by others.22 Thus, in many cases, the only way to
prevent public disclosure may be to prevent the innovation from occurring at all.

Across-the-board grant-backs might be a solution to at least the first of these two
problems, but even if sellers could perfectly draft and enforce such arrangements, we have to
ask whether it is beneficial for the law to assist them.23 That is, while it is perfectly rational
for sellers to restrain competition, we have to ask whether the law should facilitate such
restraints. When such grant-back arrangements are a product of negotiated agreements
between sophisticated parties, the cost and difficulty of negotiating such arrangements places
an inherent limit on their proliferation. But when such arrangements become too cheap and
easy to enforce, their very ubiquity makes them much more problematic.24 And although
antitrust law may help to police some of these arrangements,25 a negative impact on user
innovation can occur even in cases where there are no traditional antitrust concerns, and so
antitrust law is not a sufficient limit on these types of governance rules.

20 These phenomena are: feelings of psychological ownership, endowment effects, noneconomic motivations
and shortsightedness. See generally, Chapter 3. These phenomena affect the behavior of both buyers and sellers.
21 In the case of user innovations that are market substitutes for the seller’s products, a rational seller would
want to use exclusive grant-backs in order to suppress the innovations.
22 See generally, Part 2, Intentional Innovation.
23 And indeed, grant-backs do invite antitrust scrutiny. Exclusive grant-backs are more problematic than non-
exclusive grant-backs, for example. See Antitrust Handbook, supra note 19 at 74-5
24 See generally, Radin, Commitment, supra note 3.
25 Antitrust Handbook, supra note 19 at 74-5. Because it frequently replaces governance rules with rules of
exclusion, antitrust law may be another body of law that reflects the law’s preference for exclusion. Thanks are
due to Ariel Katz for this insight.
Some Terminological Clarifications

Before we move on to our examination of the case studies of user innovation, some discussion of terminology is in order. As we will see, in many ways the terms ‘search’ and ‘new use’ inadequately convey the full range of innovative activities 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. In part, this is because the ‘new use’ terminology is more suited to the case of real property, where any change in asset use is literally a new use. Personal property and intellectual property are by their nature more complex. 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. This may involve using the asset to make a prototype of the innovation. 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).

Two additional clarifications 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

26 Infra note 57 and accompanying text.
27 Infra note 59 and accompanying text.
28 Infra note 63 and accompanying text.
29 Infra note 57 and accompanying text.
30 See e.g. Feist Publ’ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991): “Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity. To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, "no matter how crude, humble or obvious" it might be. Originality does not signify novelty; a work may be original even though it closely resembles other works so long as the similarity is fortuitous, not the result of copying.” (Citations removed.)
31 See e.g. Bonito Boats v. Thunder Craft Boats, 489 U.S. 141, 150-151 (1989): “The federal patent system thus embodies a carefully crafted bargain for encouraging the creation and disclosure of new, useful, and nonobvious
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.32 Most patents, for example, have no economic value at all.33 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 models34 – 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.”35 Some have gone so far as to suggest that the current focus on ‘high-tech’ innovation in policy circles may be counter-productive and even harmful.36 In other words, when it comes to economic success, information production (the generation of incremental improvements, new uses, and new ideas) may often be more important than the production of intellectual property.

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 it has sometimes been suggested that ‘innovation’ should only be used to refer to commercialization, and ‘invention’ should only be used to refer to the act of creation,37 in this thesis the terms innovation and invention are used interchangeably. This accords with the term’s usage in the growing literature on ‘user innovation’ which studies, analyses and advances in technology and design in return for the exclusive right to practice the invention for a period of years.” (Emphasis added.)

34 For an example of how the law’s refusal to allow use restrictions helped to transform a business model, and even an industry, see ‘The ‘Mere’ Right to Rent/Resell Fosters User Innovation’, infra Chapter 5.
35 David Jacobson, Low Tech and R&D Spending: Problems in Policy for Innovation (April 2005) (Manuscript, on file with author.)
37 See e.g. Jan Fagerberg, Innovation: A Guide to the Literature, in The Oxford Handbook of Innovation 1, 21 (Jan Fagerberg, David Mowery, & Richard Nelson eds., 2005): “A consistent use of the terms inventions and innovation might be to reserve these for the first time occurrence of the idea/concept and commercialization, respectively.”
documents the innovative activities of users of assets. In this literature, a ‘user’ is defined as anyone (including a firm, manufacturer, individual consumer or self-employed professional) who expects to extract value from an asset by using it, rather than by selling it.38 When a user innovates, therefore, a new asset is often created, but it is not initially 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,39 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. The difference between users (which can be individuals or firms) and manufacturers is best explained by Eric von Hippel, who was the first to formulate the distinction:

“Users… are firms or individual consumers that expect to benefit from using a product or service. In contrast, manufacturers expect to benefit from selling a product or a service. A firm or an individual can have different relationships to different products or innovations. For example, Boeing is a manufacturer of airplanes, but is also a user of machine tools.”40

In this thesis, the term ‘user innovation’ is similar to, but somewhat broader than, the definition used in the user innovation literature. It would be more accurate to refer to the type of innovation that is discussed in this thesis as ‘possessor innovation.’ That is to say, whether the motivation for innovation is personal use or profit, it is critical that users be able to freely use and manipulate assets. As von Hippel’s explanation makes clear, the line between users and manufacturers is not airtight: some individuals who innovate only for use later become manufacturers because only their level of expertise can bring the product to market.41 Despite this slight difference in focus from von Hippel’s work, we adopt his term ‘user innovation’ because it most accurately reflects the process by which innovation emerges (from use), regardless of the motivation for the innovation. This theory’s focus on the innovation that emerges organically out of asset use is important because it makes it impossible to contract for the results of user innovation ex ante.

38 VON HIPPEL, supra note 1 at 3.
39 This often happens in the case of sporting goods, for example. See VON HIPPEL, supra note 1 at 127.
40 VON HIPPEL, supra note 1 at 3.
41 See note 63 and accompanying text.
One final note: it is indisputably true that, when courts replaces a governance rule with a rule of exclusion, they do not make an inquiry into whether the new use to be made of a burdened asset is of greater value than that which existed under the servitude. Courts do not even ask whether any new use has, in fact, been made or found. Nonetheless, whenever a governance rule is replaced by a rule of exclusion, the freedom that results can, over time, generate four possible outcomes: 1) there is a successful search for a new use (intentional innovation); 2) there is no directed search, but freedom of use allows a fortuitous accident which reveals a new use (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 downstream property owner appears to receive 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. 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, regardless of whether any new use is found, or any search is undertaken.

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 various types of user innovation occur.

Part 2  INTENTIONAL INNOVATION

Most of the cases that follow, because they involve the use and modification of personal property, do not involve the actual violation of a governance rule (though some of them do). The reason for this is straightforward: monitoring and enforcing restrictions on the use of personal property has historically been too difficult and too expensive. Nonetheless, in thinking about the significance of what follows it is important to ask ourselves a series of

42 The doctrine of changed circumstances might, however, be considered to be an example of this type of calculus.
43 We will discuss why this is so later in this chapter. Infra Part 5, Use, Possession and the Emergent Nature of Innovation.
44 There may be another reason as well. Because the law has generally reacted with hostility to restrictions on use, manufacturers may be aware of this, and may simply avoid what is seen to be a futile strategy.
‘what ifs’: What if permission had been required for the new use or asset modification in question - would the innovation still have occurred? And if it had occurred, would it have been disclosed, or kept secret? More importantly, what if the law were to become less hostile to use restrictions and other governance rules? Furthermore, what if the cost of implementing such restrictions was to become cheap enough to make them much more common? We will discuss some of the legal and technological developments that are leading us in this direction in Chapter 5. For now, the most important lesson to take from the examples set out below is that freely using assets is an important step in the innovative process. Thus, anything which limits use (either directly or indirectly), can adversely impact the level of innovation. Such limits should therefore be viewed with the utmost suspicion. One final caveat: some of the evidence that follows is taken from the non-user context, as innovation has been more thoroughly documented in that arena. However, to the extent that it is used, it illustrates some general points about the process of innovation, points which apply with even greater force to the user innovation case.

**Problem Finding and Problem Solving Contrasted**

That a possessor would interact differently with an asset than a non-possessor is an observation so obvious that it borders on triteness. However, the relative advantage that possessors have in interacting with assets is of critical import, because *from asset use comes innovation*. The innovation that occurs as a natural by-product of asset use will be divided into two rough categories: problem finding and problem solving. 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.\(^{46}\) Problem finding and problem solving can occur almost simultaneously, or they can be separated 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.\(^{47}\) 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).\(^{48}\)

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.\(^{49}\) 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 of 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.

---

\(^{46}\) See WALTER GRATZER, EUREKAS AND EUPHORIAS: THE OXFORD BOOK OF SCIENTIFIC ANECDOTES 104 (2002).

\(^{47}\) Sawyer, supra note 45 at 473-474.

\(^{48}\) For a brief description of this history, see JAMES W. BENSON, TIME AND TIME-TELLERS 46-52 (Harvard 1902).

\(^{49}\) For the details of the Post-It note story see Sawyer, supra note 45 at 474.
In this case, the ‘problem’ (flagging pages in a document) was not found until roughly five years after the ‘solution’ (the unusual adhesive).

In the sections that follow, we will examine instances of innovation in areas such as high-performance windsurfing, mountain biking, surgical equipment and computer software. 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 (in part, this is what makes them identifiable to researchers). 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 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).

Clearly, these stories of user innovation only tell us about the successful innovations, rather than about the failed innovations or the cases in which users do not innovate at all, and so there is always a chance that they suffer from selection or survivorship bias. That is, because we see the successes and not the failures, we overestimate the prevalence of the successes. However, there are several reasons why the presence of selection bias does not weaken the importance of users to the innovative process. First, even if many or most of users do not innovate - and it is assumed that most do not – those who do innovate will still produce qualitatively significant advances, because users have advantages in producing innovations that emerge from asset use. Second, there is evidence that users in fact do produce a quantitatively significant amount of innovation, and so there may be no selection bias at work in this data. Third, although we are focusing on user innovation, we are not saying that nonusers do not innovate; we are simply saying that it is advantageous if users are...

---

50 See e.g. PAUL ORMEROD, WHY MOST THINGS FAIL 12 (2005): "The tendency to overemphasize successes, and to rationalize them ex post is chronically endemic amongst business historians. . . . It is failure rather than success which is the distinguishing feature of corporate life”; and LEONARD MLODINOW, THE DRUNKARD’S WALK 178-181 (2008), discussing the “hot-hand fallacy,” a common misperception that repeated success resulted from specific actions rather than chance, and saying that “among a large group . . . it would be very odd if one of them didn’t experience a long streak of successes or failures.”

51 See infra notes 56-96 and accompanying text.

52 See infra notes 97-113 and accompanying text.
also given the freedom to innovate. Fourth, it is inherently impossible to predict who will innovate and who will not, and so the legal preference for exclusion performs an important role in ensuring that those who can innovate via asset use will in fact be able to do so.53

It is no accident that many (though not all) of the examples below are from the fields of individual sports and computer software/programming. Both of these fields share 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. Under such conditions, users have advantages in both problem finding and problem solving. It has long been recognized, for example, that the movement from hard-wired integrated circuits (which had dedicated functions that were determined by the physical configuration of the wires on the chip), to the programmable microprocessor (which had open-ended functionality that could then be determined by directions coded into software), marked a watershed in the proliferation of computerized devices.54 The move from hardware coding to software coding, along with the general purpose nature of the microprocessor, meant that “[i]ts functions could be defined to a greater extent by the users themselves who could program the chip according to their needs.”55 However, the same principle applies to any asset that can be manipulated by a user, whether that asset is a mere chattel (such as a bike) or a high-tech device. The important thing, from the point of view of innovation, is that the user be able to use and modify the asset.

Even though we have seen that problem finding and problem solving are not necessarily consecutive events, we will begin our examination of user innovation case studies by discussing problem finding, and then move on to problem solving.

A) PROBLEM FINDING

There are many examples of asset use revealing problems among users – we will review just a few of them here. The case studies on user innovation reveal two salient facts about users and problem finding. First, that new problems are often revealed during the

53 See infra notes 156-160 and accompanying text.
54 ELHANAN HELPMAN, GENERAL PURPOSE TECHNOLOGIES AND ECONOMIC GROWTH 266 (MIT 1998).
55 Id.
ordinary course of using an asset; and second, that users have highly individualized patterns of use, which leads them to experience a unique set of needs (and therefore to find a unique set of problems). Take mountain biking as an example: different mountain bikers have different riding styles and specializations, and they ride over very different terrains. These unique patterns of use create unique sets of problems (i.e., equipment needs), which are revealed through the use of standard equipment. As a result, user innovation among mountain bikers is highly diverse. In one study “at most 10 of 43 innovations had functionality similar to that of another sample member.”56 Here are two examples of how equipment use can reveal problems, and how equipment modification can play a role in user innovation, in the sport of mountain biking:

“I do back-country touring and needed a way to easily lift and carry a fully loaded mountain bike on the sides of steep hills and mountains and dangle it over cliffs as I climbed. Solution devised: I modified the top tube and the top of my seat post to provide secure attachment points for a carrying strap, then I modified a very plush and durable mountaineering sling to serve as the over-shoulder strap.

Because the strap sits up high, I only need to bend my knees a little bit to lift the bike onto my shoulders, yet it is just high enough to keep the front wheel from hitting when I am climbing a steep hill. Eventually, I came up with a quick-release lateral strap to keep the main strap from sliding off my shoulder, but it will easily break away if I fall or land in a fast river and need to ditch my bike.

When riding on ice, my bike has no traction and I slip and fall. Solution devised: I increased the traction of my tires by getting some metal studs used by the auto industry for winter tires. Then I selected some mountain biking tires with large blocks of rubber in the tread pattern, drilled a hole in the center of each block and inserted a stud in each hole.”57

Interestingly enough, despite the highly individualized nature of their needs, many of the innovations that users eventually create eventually end up being useful to a significant number of other users. This pattern might seem somewhat unusual, but it accords with Eric

56 VON HIPPEL, supra note 1 at 35.
von Hippel’s analysis of ‘lead users.’ Although users have unique patterns of use which give rise to unique needs, in many cases, those who innovate are simply ‘ahead of the curve’ in terms of asset use. That is, other users sometimes ‘catch up’ and change their patterns of use, either as a natural consequence of increasing sophistication over time, or by deliberately adopting the use patterns of the lead users. One common example of this type of evolution occurs when a sports enthusiast observes the ‘new moves’ of the best players, and deliberately seeks to imitate them. At times, this might only involve the imitation of technique; at other times, it might also involve the adoption of an innovative technology. A good example of the latter situation can be seen in the development of high-performance windsurfing:

“There was new enthusiasm for jumping and we were all trying to outdo each other by jumping higher and higher. The problem was that … the riders flew off in mid-air because there was no way to keep the board with you – and as a result you hurt your feet, your legs, and the board.

Then I remembered the ‘Chip,’ a small experimental board we had built with footstraps, and thought ‘It’s dumb not to use this for jumping.’ That’s when I first started jumping with footstraps and discovering controlled flight. I could go so much faster than I ever thought and when you hit a wave it was like a motorcycle rider hitting a ramp; you just flew into the air. All of a sudden not only could you fly in to the air, but you could land the thing, and not only that, but you could change direction in the air!

The whole sport of high-performance windsurfing really started from that. As soon as I did it, there were about ten of us who sailed all the time together and within one or two days there were various boards out there that had footstraps of various kinds on them … . It just kind of snowballed from there.”

This scenario illustrates several features that are common in user innovation. Once again, we see that a solution (the ‘Chip’) can be created before a problem is identified (falling off the board). We also see the repeated use of a manufacturer-produced asset revealing shortcomings in its design. A user at the forefront of asset use innovates in a fairly simple, low-tech manner that turns out to be extremely valuable. It is then adopted by other users, who modify their existing manufacturer-produced assets in order to duplicate the

58 See generally VON HIPPEL, supra note 1 at 19-31.
innovation. In this case, the user innovation not only ‘improved’ the state of innovation in its field of use, it literally created a new sport. Over time, the user-innovators of the new boards became manufacturers themselves, as others entered the field and requested that the user-innovators make boards for them. This is common in the area of sports equipment. In other areas, users do not become manufacturers themselves, but do coordinate with manufacturers who bring the innovation to market. One study in Germany found that: “22 percent [of surgeons surveyed] reported developing or improving some item(s) of medical equipment for use in their own practices.” Of these innovations, 48 percent were eventually marketed by mainstream medical manufacturers. And in other cases, such as mountain biking, a mix of both user-manufacturers and mainstream manufacturers evolves over time. Users who originate the new sport first build the products they need from scratch, and then move on to building bikes for others. Finally, traditional manufacturers gradually enter the field, and co-exist with the user-originated firms.

As you read the following excerpt, consider that manufacturers sometimes sell their tangible products with labels that purport to restrict the ways in which they are used or modified. For example, vacuum tubes have been sold with labels stating that they were ‘licensed’ for use in car chargers, but not radios. What if the manufacturers of the motorcycle parts in the following scenario had printed field-of-use restrictions on the parts that they produced?

“Mountain biking began in the early 1970s when some young cyclists started to use their bicycles off-road. Existing commercial bikes were not suited to this type of rough use, so early users put together their own bikes. They used strong bike frames, balloon tires, and powerful drum brakes designed for motorcycles…

Commercial manufacturing of mountain bikes began in 1975, when some of the early users…began to also build bikes for others. A tiny cottage industry developed, and by 1976 a half-dozen small assemblies existed in Marin Country, California. … Major bike manufacturers then followed and started to produce mountain bikes and sell them at regular bike shops across the United States.”

---

60 VON HIPPEL, supra note 1 at 30.
61 Id.
63 VON HIPPEL, supra note 1 at 72-73.
Further evidence of the advantages that users have in problem finding is provided by the fact that users and manufacturers create different types of innovations. When researchers compared the innovations of users and manufacturers in the area of scientific equipment, for example, they found an interesting pattern: users tended to develop novel functions, while manufacturers tended to improve existing functions.\(^{64}\) What is the relevance of this pattern? It illustrates that asset use provides an important (and irreplaceable) source of new problems (i.e., it creates advantages in problem finding). That is to say, the need for novel functions is revealed by asset use, and asset use itself creates new needs. Unique patterns of use across different users will therefore reveal new needs, which are then followed by a search for solutions (novel functions).

This leads us to the first assertion about user innovation:

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

B) PROBLEM SOLVING

In the conventional model of innovation, a manufacturer comes up with a product idea in response to a consumer want or need (the problem), develops it into a finished product (the solution), and places it on the market.\(^{65}\) 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. 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.

\(^{64}\) Id. at 71.
\(^{65}\) Id. at 2.
i) Asset Use Provides a Trial and Error Process

The process of problem solving is, in all fields, a trial and error process in which an innovator responds to a problem by seeking out a tentative solution based on his knowledge and skills, building a prototype or model, and testing the performance of that prototype or model under conditions that simulate product use. The information from that trial (the error) is then used to either modify the tentative solution, or to seek a wholly new solution in a different direction. This process is repeated, and the solution is sequentially refined. It is easy to see why users would have an advantage in problem solving around their own needs: repeated asset use not only naturally reveals problems; it also provides an organic testing ground for attempted solutions. If a skateboarder modifies his board to improve his ability to perform an unusual move, for example, but ends up worsening his overall performance, he not only knows that his initial solution did not work; he has also gathered information about how the modification impacted his performance, and this data can be used to direct the path of the next attempted solution. Users thus have a relatively low-cost means of testing and improving their innovations.

In some ways, it seems inapt to talk about the ‘cost’ of user innovation, at least when it occurs as an organic by-product of asset use. After all, many users (such as sports enthusiasts) are engaging in asset use for fun and enjoyment, and the purpose of their innovative activity is to increase that fun and enjoyment. But users have cost advantages (i.e., relative to manufacturers) in problem solving in other ways as well. Studies show that the majority of user-innovators use skills and knowledge that they already have to develop their innovations, thus eliminating the need to invest in learning something new. In one study, more than 50 percent of users said that they used information they had acquired from the activity itself (e.g., from engaging in their preferred sporting activity or from another hobby) to develop their innovations; while almost 48 percent said that they used information from their professional background. Thus we can see that asset use not only provides an ideal


67 VON HIPPEL, supra note 1 at 74.
way to test innovations; it is itself an important source of what Arrow called ‘learning by doing,’ and what Rosenberg called ‘learning by using.’

**ii) Different Users Will Perceive Different Solutions**

There are various reasons why we all have a unique way of perceiving the world. At the most basic level, we start off with different physical skills and innate talents. Some of us have better eyesight or better dexterity. Some of us are better at analyzing spatial relationships, while others are better at language. To this we add the different skill sets and knowledge that we accumulate and develop over time. But below this there is something even more fundamental: we all have different internal representations of the world, and this means that each of us literally sees the world differently. These internal representations are called perspectives: “a perspective is a map from reality to an internal language such that each distinct object, situation, problem, or event gets mapped to a unique word.” The result of this internal translation process is that users are irreplaceable: one user cannot be replaced with another. Why is it so important to bring different perspectives to bear on the problem solving process? For the skeptical, it might be best to quote from an economist and mathematician: “[because] the right perspective can make a problem easy. …[M]ost scientific breakthroughs and business innovations involve a person seeing a problem or situation differently.” This phenomenon is what makes much of the computer software produced by open source communities so technically advanced: the eyes of many users not only uncover problems that go unnoticed by others; each pair of eyes literally sees the problems in different ways. Von Hippel has analyzed this phenomenon in terms of users having different ‘toolkits,’ which allow different users to match their skill set to the appropriate problem. Having the right skill set and experience doubtless plays an important role in creating the advantages of user innovation. However, the assertion here is that users have advantages in solving problems for two reasons: 1) because different users have

---

71 Id. at 25.
72 See generally Von Hippel, supra note 1 at 16, 94, 100-01, 128, 147-64.
different skill sets; and 2) because users have different perspectives which literally means that the problems are different when viewed by different users.73 Together, these unique skill sets and perspectives will be referred to as a user’s ‘perception’.74

There is a related reason why users’ ability to freely interact with the assets that they possess is a key step in the cultivation of user innovation: experience has shown that we cannot predict who will innovate, how they will innovate, or in what direction they will innovate. One illustration of this is that non-experts often solve problems that befuddle experts: an appropriate ‘toolkit’ (unless we include unique perspectives in our definition of ‘toolkit’) does not guarantee a solution. Indeed, many of the famous people that we think of as ‘experts’ actually had no training in the fields in which they innovated. Darwin, for example, had no formal training in zoology or botany when he developed the theory of natural selection which changed the course of the biological sciences.75 He was actually a self-taught geologist when he took part in the famous voyage of the H.M.S. Beagle. But perhaps what is more surprising is the fact that John Gould, the ornithologist who analyzed the information that Darwin had gathered about the birds of the Galapagos, did not recognize the importance of the data. Darwin thought that the birds he had documented, which had such different beaks, were all different birds. It was Gould who informed him that they were actually all finches. But despite Gould’s knowledge of ornithology and taxonomy, it was Darwin who realized the significance of the differentiation among the finches – that it represented a gradual adaptation of the birds to their unique island environments.

The surprising tendency for non-experts and professionals to make major advances completely outside their field of knowledge and training has been colloquially termed ‘the Medici effect.’76 And this trend is not limited to science: it occurs in all areas of innovation, big and small, commercial and noncommercial.77 In fact, this phenomenon is so important that businesses are beginning to take advantage of it to solve their problems. Using the Internet, some companies are advertising their most intransigent innovation problems to the world, in the hopes that more eyes might see a solution that in-house experts had missed.

73 PAGE, DIFFERENCE, supra note 70 at 153-154.
74 Scott Page discusses what I refer to as ‘perception’ in terms of an individual’s perspective, heuristics and toolbox. See generally PAGE, DIFFERENCE, supra note 70.
76 See generally, id.
77 On the tendency of ‘outsiders’ to innovate, see e.g. ANDREW S. GROVE, ONLY THE PARANOID SURVIVE (Random House 1996).
This practice is known as ‘broadcasting’, and it has produced successful solutions in a significant number of cases: the most comprehensive study of the strategy to date has found that one-third of the problems advertised through this method were solved quickly and cost-effectively. More importantly, they were solved by individuals completely outside of the field in which the problem occurred. Once again, we see that being able to conceive of problems in a different light, to literally see a different problem, is frequently the first step in a successful process of innovation. Although the phenomenon of broadcasting also shows us that both nonusers and users can innovate, this does not undercut the importance of user innovation. Both users and nonusers can and will innovate, if they are given opportunities to do so. The key point in this section is that we cannot predict who will give us the right answer (who will successfully innovate). This inability to match assets to innovators on an ex ante basis means that the best that we can do is to allow a variety of individuals to try. In other words, the legal preference for exclusion operates as an ad hoc matching process: by giving possessors of assets full and free use of those assets, the law maximizes the chances that an individual will be able to innovate when he happens to be paired with the ‘right’ asset.

iii) 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. Problems can occur when innovators prefer solutions that require them to violate or ignore

79 Martha Lagace, Open Source Science: A New Model for Innovation: Q&A with Karim Lakhani, HARVARD BUSINESS SCHOOL WORKING KNOWLEDGE, Nov. 20, 2006, available at http://hbswk.hbs.edu/item/5544.html. “The winning solution was proposed by a scientist from Finland who did not work in this field.”
80 Id. “Innovations happen at the intersection of disciplines. … The best toxicologist within the firm had a look at the results and couldn’t understand them … Then they finally posted in on InnoCentive. A protein crystallographer looked at the problem and basically gave an off-the-shelf solution.”
81 See infra notes 157-160 and accompanying text.
governance rules that manufacturers seek to impose on assets - this may hamper implementation of an innovator’s preferred solution. A simple example of this occurs with cartridges for ink-jet printers. Some users, when faced with novel printing problems, came up with solutions that involved refilling ink-jet cartridges with inks not sold by the manufacturer. Some users, for example, wanted to ‘print’ on cakes, so they refilled the cartridges with food dye. Others wanted to use their ink-jet printers to print photo-quality pictures, so they refilled the cartridges with specialized ink.

These certainly weren’t the only options available to these users: the user who wanted photo-quality printing could simply have purchased another type of printer, for example. Nonetheless, refilling the cartridges is often the users’ preferred solution. But many manufacturers of ink-jet printers do not want consumers to be able to refill their cartridges, because they sell their printers at low margins, and their cartridges at high margins. The logic behind this strategy is that a consumer only buys a single printer, but he buys many cartridges; implementing a high-margin strategy on the cartridges thus provides a higher ongoing revenue stream. Of course, this business model is threatened if consumers can reuse (i.e., refill) the ink cartridges, so manufacturers try to restrict the ability of users to refill the cartridges. They use two primary methods to implement these restrictions: contracts and technological controls. These barriers impose a governance rule on the cartridge (i.e., separate some of the use rights from the right of possession), because a user can no longer reuse the cartridge however he likes.

The preferences of innovators can obviously impact the solution set that is generated and revealed to the public, and so it is advantageous to have multiple innovators pursuing and revealing multiple solutions. But even when this is the case, the preferences of innovators can sometimes impact the direction and the pace of innovation. The Wright brothers provide us with a famous example in which innovator preferences had an impact on both the direction

---

82 VON HIPPEL, supra note 1 at 117.
83 Id.
84 When a contract is used to implement a use restriction, the question for the courts is whether to enforce that contract. But even when the method of implementation is technological, there is often a legal question involved. This is because technological restrictions are usually implemented via software code, which is protected under the Copyright Act, and because these technological restrictions can usually be circumvented in some way. The Copyright Act contains provisions which make the circumvention of technological restrictions unlawful in many circumstances. For more on this issue, see infra Chapter 5, Part 2, The Future of The Numerus Clausus - The Challenge of Technology.
and pace of intentional innovation.\textsuperscript{85} The Wright brother example is particularly useful because it is meticulously documented, and because it provides persuasive evidence of both the pervasiveness of diverse innovator preferences, and their relative impact.\textsuperscript{86} 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.\textsuperscript{87} 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); therefore, when otherwise savvy innovators stubbornly insisted on pursuing or adopting suboptimal solutions, the impact of preferences is made particularly apparent.

The Wrights were pioneers in the field of aviation, being the first to achieve heavier-than-air manned flight.\textsuperscript{88} They set aviation records, and held the most important patent in the field of aeronautics: the wing-warping patent.\textsuperscript{89} But if we take a 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).\textsuperscript{90} 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

\textsuperscript{85} The Wright brothers were not users in the sense that is used in the literature on user innovation (they were innovating primarily for manufacture, rather than just for their own use). However, as used in this work, the term user is broader, being roughly synonymous with a possessor who uses an asset. In addition, it should be noted that the individual inventor, driven by both the enjoyment and intellectual challenge of inventing and (sometimes) the prospect of inventing a product that some manufacturer might want to make, straddles the line between amateur users and manufacturers. Indeed, scientists also straddle this line. Thus, it is appropriate to look at the innovative activities of inventors and scientists along with amateur users.

\textsuperscript{86} If preferences - so infused with 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 amateur user-innovators to an even greater degree.

\textsuperscript{87} See generally Fred Howard, Wilbur and Orville: A Biography of the Wright Brothers 133-47 (Dover 1998) [Howard].

\textsuperscript{88} Id.

\textsuperscript{89} Id.

\textsuperscript{90} Id.
apparatus were no longer needed. But even after the Wrights became aware of this innovation, and its rapid adoption, 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.”’ 91

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, 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, because he felt that this would obstruct the view of scouting planes.92 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.93 But no tractor planes were produced by the Wright Company until after Orville left in 1915.94

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 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.95

Orville’s responses can be attributed to a variety of factors (other than to cognitive inability). In some cases, he 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 weighed certain considerations differently than others (as in the case of the

91 SETH SHULMAN, UNLOCKING THE SKY 128 (Perennial 2003)(2002). This is especially surprising when we consider that the Wright brothers used to own their own bicycle shop.
92 HOWARD, supra note 87 at 392.
93 Id.
94 Id.
95 Id. at 133-47.
pusher plane). And in other cases, he might have simply been emotionally attached to his own solution.

This pattern of stubbornness in the face of objective evidence 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, innovations and scientific advances often owe more to the relentless pursuit of idiosyncratic beliefs and preferences (contrary to all available evidence), than to the pursuit of logical solutions. 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 for the downside of innovator intransigence is, of course, to have multiple innovators finding and promoting alternative solutions.

This leads us to the second assertion about user innovation:

2. Users of assets will have an advantage in problem solving because:

   i) asset use provides a low-cost, repeated, trial and error process;

   ii) differential cognition across users will lead users to perceive different sets of possible solutions to problems; and

   iii) 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.

   All of these characteristics increase the likelihood of locating a successful solution (i.e., generating a useful innovation).

Part 3 THE ECONOMIC IMPORTANCE OF USER INNOVATION

The case studies clearly show that users are quite active in innovating around the assets that they use and possess. But is all of this innovative activity on the part of users economically important? There are several reasons to think that it is. First, user innovation is important because it provides individual users with isolated solutions to their individual

96 On the need for scientists to go beyond the available evidence, see generally DAVID FAUST, THE LIMITS OF SCIENTIFIC REASONING (University of Minnesota Press 1984).
problems, and moves specific assets to higher-value uses. Even 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. Businesses are becoming ever more aware that they can adopt and build on the actual innovations of their customers (rather than just their ‘product suggestions’),97 and some niche industries are built almost entirely on the commercial adoption of such innovation (this was the case with mountain biking equipment, for example). 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.”98 (Emphasis added.) The evidence about the commercial significance of user innovation is even more compelling when we include data on user firms that develop new products or techniques for use in-house.99

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., a plumber or electrician modifies a tool to make his job easier). However, whether a user is an individual or a member of a firm does not matter for our purposes. Whether innovation is produced by an individual tinkering at home in his garage, or by a factory worker on an assembly line, such activity usually involves the modification or modified use of pre-existing assets, which makes it important to unify use and possession across all users.

Second, there is evidence that individual users often produce innovations that are more advanced than those created by manufacturers. In the case of computer software, for example, some of the most successful (and technically sophisticated) products were developed by open source communities. The ‘Apache’ web sever software, for example, is

97 It should be noted that building on customer innovations and responding to the product requests are two very different things. See VON HIPPTEL, supra note 1 at 144. Thus, evidence that companies who are too attentive to customer demands miss out on major innovations - see e.g. CLAYTON CHRISTENSEN, THE INNOVATOR’S DILEMMA 59 (Harvard University Press 1997) – do not apply to the adoption of customer innovations.
98 VON HIPPTEL, supra note 1 at 2.
99 For examples, see JOHN ENOS, PETROLEUM PROGRESS AND PROFITS (MIT Press 1962); C. Freeman, Chemical Process Plant: Innovation and the World Market, 45 NAT’L INST. ECON. REV. 29 (1968); and Keith Pavitt, Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory, 13 RESEARCH POLICY 343 (19984).
used by more than 60 percent of the world’s websites, and it is considered to be superior to commercially developed web server products. And individual users can have even more innovative impact than innovation communities. The development of the Internet, for example, is replete with examples of individuals creating the completely unpredicted ‘killer app.’ But computer geeks are not the only ones who produce innovations of real and lasting market value. In some fields, user innovation is literally responsible for the majority of significant advances: “the most commercially important equipment innovations in four sporting fields tended to be developed by individual users.” Although these extreme sports may appear to be ‘niche’ markets, they are nonetheless quite lucrative: “In 2000, about $58 billion (65 percent) of total retail sales in the US bicycle market were generated in the mountain bike category.”

Third, even when user innovations are ‘low-tech’ or relatively simple, they can still pack an economic punch. It is true that many user innovations are incremental or mundane character, but this need not cause us to question their economic relevance. ‘Small’ advances can in fact be the most economically important: “most innovations from any source [i.e., industry and users] are minor… [but] minor innovations are cumulatively responsible for much or most technical progress.”

Fourth, the data on user innovation, substantial and growing though it is, does not give us the full picture of its extent and impact. To get a better sense of the true extent of user innovation, we also need to consider the literature on general purpose and generative technologies. A general purpose technology is “a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects.” General purposes technologies include everything from bronze and the steam engine, to the laser and

100 VON HIPPEL, supra note 1 at 101.
102 VON HIPPEL, supra note 1 at 22 (citing evidence gathered by Shah, supra note 59).
103 Id. at 73.
104 VON HIPPEL, supra note 1 at 21.
105 Some of the innovations measured and discussed in the literature on user innovation will also be discussed and measured in the literatures on general purpose and generative technologies; but the overlap will not be complete. Thus, although we do not suggest that these literatures are strictly additive, they nonetheless all warrant inclusion if we are to get a comprehensive view of the scope of user innovation.
the computer. General purpose technologies are key elements to economic growth over time, because they are susceptible to a variety of unforeseen new uses, new applications and new modifications on the part of a multitude of developers and users.

Generative technologies (and systems) are similar to general purpose technologies, but these categories do not completely overlap: “Generativity is a system’s capacity to produce unanticipated change through unfiltered contributions from broad and varied audiences.” The characteristics of any generative technology or system are: leverage, adaptability, ease of mastery, accessibility and transferability. Considering these factors, we can see that a computer is both a general purpose and a generative technology, while a laser is not. With generative technologies and systems, the driving force of innovation is the amateur, while general purpose technologies are developed by a mix of amateurs and experts (depending on the sophistication of each particular technology). Despite their differences, both general purpose and generative technologies contribute greatly to economic growth because they can be developed in a variety of directions by a variety of users. In other words, both general purpose and generative technologies make the most efficient use of human capital inputs.

Finally, alongside the demonstrated success of user innovation, we find the failure of a significant number of manufacturer innovations: “most new products developed and introduced to the market by manufacturers are commercial failures.” For all of these reasons, it is accurate to say that user innovation is an important, and irreplaceable, source of economic value. Indeed, the lead empirical researcher in this area has concluded that: “an innovation system where user innovation is present is welfare superior to one where it is not.”

107 Id.
108 Id. at 1.
110 Id. at 71-73.
111 The importance of this is discussed infra notes 161-179 and accompanying text.
112 VON HIPPEL, supra note 1 at 107.
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: ‘Post-It’ notes, 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.\(^{114}\) 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: 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.\(^{115}\) 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).\(^{116}\) 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.\(^{117}\) 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.\(^{118}\) Even errant lab members have 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.\(^{119}\) 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

\(^{114}\) See generally ROYSTON M. ROBERTS, SERENDIPITY: ACCIDENTAL DISCOVERIES IN SCIENCE (1989) [Roberts].

\(^{115}\) Id. at 71.

\(^{116}\) Id.

\(^{117}\) Id. at 225.

\(^{118}\) Id. at 177.

\(^{119}\) For a retelling of this story, see id. at 172-74.
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. 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.

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 and 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. 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

---

120 Later other useful polyesters were found, and were used in a variety of products, including textiles. Id.
121 Roberts, supra note 114 at 171: “This accident turned the failure into the enormous success advertised at the 1939 New York World’s Fair as ‘Nylon, the Synthetic Silk Made from Coal, Air, and Water!’”.
122 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. Id.
123 Sawyer, supra note 45 at 469-71.
‘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” 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 four features of interest to us: 1) the innovation was accidental; 2) the innovation was user driven; 3) the innovation was successfully adopted by a manufacturer; and 4) the innovation was economically important.

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.125

124 Sawyer, supra note 45 at 470.
125 These are adopted from the accident “intensity scale,” described in Robert D. Austin and Lee Devin, Accident, Intention, and Expectation in Innovation Process 4-10 (Working Paper No. 06-026, 2008).
i) 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\(^{126}\) to Salvador Luria,\(^{127}\) 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.”\(^{128}\)

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.\(^{129}\) 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.

---

\(^{126}\) See e.g. Ernst Mach, On the part played by accident in invention and discovery, The Monist, Jan. 1896, at 161-175.


\(^{129}\) Donald Campbell, Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes, 67 Psychological Review 380, 397-98 (1960).
which is discontinuous with the previous stock of knowledge. This is the only way we can account for “repeated ‘breakouts’ from the limits of available wisdom.”130 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.131

Although Campbell was initially concerned with the blind variation that occurs in the thought processes of any one individual,132 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.133 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.”134 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’135,

---

130 Id. at 380.
131 Discontinuity (without antecedent) does not imply that the innovation be ‘radical’ in terms of impact.
132 Though Campbell did mention the parallels between his theory of creative thought and Mach’s thinking on innovation.
133 Actually, Campbell’s blind variation-selective retention model has been applied to a multitude of evolutionary processes, such as organizational science and cultural evolution.
134 W.B. Cannon, The Role of Chance in Discovery, 50 SCIENTIFIC MONTHLY 204-09, 207 (1940) [Cannon].
135 The famous quote from Pasteur is often recorded as: “Chance favors the prepared mind.” It is actually: “In the field of observation, chance only favors the prepared mind.” This translation comes from Hugo Kubinyi, Chance Favors the Prepared Mind—From Serendipity to Rational Drug Design, 19 J. RECEPTOR & SIGNAL TRANSDUCTION 15, 18-19 (1999).
‘intelligent observer’\textsuperscript{136} and ‘sagacity’\textsuperscript{137} 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.”\textsuperscript{138}

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 capacious definition of

\footnotesize{\textsuperscript{136} Cannon, supra note 134 at 207.\
\textsuperscript{137} Id. at 208: “There was an instance of serendipity—a discovery which I was ‘not in quest of,’ a disclosure which called for the application of ‘sagacity,’ to use Horace Walpole’s expression” (emphasis added).\
\textsuperscript{138} 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. GRATZER, supra note 127 at 182.}
sagacity that goes beyond classical training, and embracing 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.\(^{139}\) (We might expect few contributions to theoretical physics by the uneducated, but 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,\(^{140}\) and remains the best modern cure for malaria – which is still one of the most lethal diseases worldwide.\(^{141}\) 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.\(^{142}\) 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.

\textit{ii) 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 individuals 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 accidents do not simply speed up the pace of scientific discovery, they also lead to the creation of wholly new products. Someone would have

\(^{139}\) Campbell,\textit{ supra} note 129 at 397-98.
\(^{140}\) This is the first documented usage. See ROBERTS,\textit{ supra} note 114 at 6-10.
\(^{141}\) 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.\textit{ Id.}
\(^{142}\) \textit{Id.}
eventually discovered X-rays, for example, even though Röntgen discovered them much faster thanks to an accident.\textsuperscript{143} But would someone have inevitably created Velcro, whose invention stemmed from George deMestral’s frustration with removing cockleburs from his coat?\textsuperscript{144} How about floating soap (Ivory Soap), which was created when a worker mistakenly left a stirrer on during lunch?\textsuperscript{145} 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.\textsuperscript{146} 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.\textsuperscript{147} 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

\begin{footnotesize}
\textsuperscript{143} Id. at 139.
\textsuperscript{144} ROBERTS, supra 114 at 220.
\textsuperscript{145} Id. at 220.
\textsuperscript{146} See Campbell, supra note 129 at 385-386.
\textsuperscript{147} Id. at 387, quoting Ernst Mach: “Thus are to be explained the statements of Newton, Mozart, Richard Wagner, and others, when they say that thoughts, melodies, and harmonies had poured in upon them, and that they had simply retained the right ones.”
\end{footnotesize}
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 Austin and Devin’s 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, the search. 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. 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 many of the major advances in human history: “almost no discovery, no technologies of note, came from design and planning – they were just Black Swans.” A Black Swan is an event which is rare, unpredictable, and has extreme impact. For example, the discovery of

---


149 Austin and Devin, *supra* note 125 at 22, 23-30.

150 *Id.* at 24-26.

151 *Id.* at 21.

152 *Id.*

cosmic background radiation, which helped us to understand how the universe was formed, was a Black Swan: hailed as one of the twentieth century’s greatest advances, it was discovered while researchers were looking for bird poop.\textsuperscript{154} Although such ‘Black Swan’ advances can be understood in terms of a series of connected events after they occur, they could never 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 that 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.

\textbf{iii) 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 preference for exclusion 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

\textsuperscript{154} \textit{Id. at} 166.
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, exclusionary rules in property law, which appear inefficient because they apply across the board, are actually more efficient (value enhancing) in the long run – precisely because they apply 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 and 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 legal preference for rules of exclusion does.

Part 5  USE, POSSESSION AND THE EMERGENT NATURE OF INNOVATION

The evidence shows that user innovation is a significant source of value, for both users and manufacturers. But why can’t buyers and sellers contract _ex ante_ for appropriate

---

155 Austin and Lee, _supra_ note 125 at 28.
divisions of rights which will foster user innovation? In most cases of user innovation, neither the buyer nor the seller has any idea at the time of sale what innovations will be created, because they emerge as natural by-product of asset use. Indeed, in most cases neither the buyer nor the seller has any idea at the time of sale that the user will innovate around the asset, so it would be impossible for the seller to factor this fully into the price. People cannot bargain over what does not exist, over what they have not even contemplated. And if a significant amount of innovation occurs 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. This, of course, is just another way to say that different allocations of rights will give rise to different patterns of use and innovation; and the reason for this is that innovation is an emergent process.

A characteristic or event is said to be emergent if it cannot be predicted from, or broken down into, its constituent components. When properties are emergent, 1 plus 1 can equal 14, but there is no way to know this in advance (or to reverse engineer the result after the fact). The process of innovation is emergent because it cannot be predicted in advance who will innovate or how they will innovate, and because a significant amount of innovation literally emerges from asset use. The emergent nature of innovation means that individuals need to interact with an asset in an unrestricted manner, in order for unexpected innovations to (possibly) emerge at some later time. The emergent nature of innovation also makes it impossible to predict which individuals need to be paired with which assets in order to generate the optimal amount of innovation. This makes it impossible to specifically contract (or to allocate rights appropriately, which is essentially the same thing) for a significant amount of innovation in advance.

The emergent nature of innovation thus gives rise to a matching problem, for which there is no first order solution. However, unifying full use rights in an asset with its permanent possession provides a way to increase the chance that the ‘right’ asset will be paired with the ‘right’ innovator. This fortuitous matching process is much like that which

---

156 After innovation has occurred, the allocations of rights will themselves create barriers to recontracting. See Chapter 3.
157 For an exhaustive discussion of the meaning of emergence, see generally Timothy O’Connor, Emergent Properties, 31 AMER. PHIL. Q. 91 (1994).
occurs with general purpose and generative technologies. In the case of general purpose technologies, the matching arises from the inherent characteristics of the technology: their very adaptability makes them amenable to a variety of uses by a variety of users. In the case of generative technologies, the matching also arises from inherent technological characteristics, but this process is often interfered with by legal and technological restrictions on use (resulting in the production of innovation-hindering ‘tied’ technologies). Of course, unifying rights of use and possession is not a perfect way to match the right asset with the right innovator: there might be individuals who do not possess particular assets, who would in fact use them in a more innovative manner. In most cases, there is little we can do about this. But recall the phenomenon of ‘broadcasting’: this is an example of private enterprise actively seeking to remedy the mismatch of individuals to innovative problems. The *numerus clausus*, via its preference for rules of exclusion, essentially performs a similar function with respect to tangibles. However, it does so passively, via the creation of a functional equivalent to a trial and error process in which every user of an asset gets full use of that asset by legal fiat.

The use of broadcasting to find innovative solutions raises a related question: why can’t industrial enterprises simply duplicate the innovative activities of users, either via an organized, methodical program of research and development, or via a process which is more flexible – one which allows accidents, the input of various employees, and collaboration across departments? The answer is that neither use, nor users, can be duplicated by even the most cleverly designed organization. One user cannot be replaced by another, as each user represents a unique set of productive opportunities. The best a manufacturer can do is to allow users to innovate, and seek out marketable user innovations. This is because every individual has a unique stock of human capital, and human capital is the key constraint in the search for higher-value uses of assets.

---

158 *See supra* notes 105-111.
159 *See supra* notes 78-81.
160 This is a simplification, of course. In general, the law unifies use and possession with respect to chattels and intellectual property embodiments, but loosens this preference with respect to land. *See Chapter 4, Part 1, 2) The Law has a Greater Tolerance for Servitudes on Land Than for Servitudes on Chattels.*
Human capital, like other forms of capital, is a durable asset that produces income and productive outputs over long periods. 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, 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.


Tacit knowledge can be transferred, but that transfer is extremely difficult. Human capital cannot be transferred from one person to another, although knowledge can be. Education, for example, does not transfer human capital from one person to another. Instead, education transfers knowledge, which in turn develops the human capital of the learner. On the distinction between human capital and knowledge (or ideas) see Paul Romer as quoted in Ronald Bailey, Post-Scarcity Prophet, REASON, December 2001, available at http://reason.com/archives/2001/12/01/post-scarcity-prophet/singlepage: “Human capital is comparable to a thing. You have skills as a writer, for example…. That's not something that we can clone and replicate. The formula for an AIDS drug, that's something you could send over the Internet or put on paper, and then everybody in the world could have access to it. This is a hard distinction for people to get used to, because there are so many tight interactions between human capital and ideas. For example, human capital is how we make ideas. … So human capital makes ideas, and ideas help make human capital. But still, they're conceptually distinct.”

See e.g. Paul Romer, Endogenous Technological Change, 98 J. POL. ECONOMY S71, S71 (1990): “the stock of human capital determines the rate of growth.”
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 combines and builds on a number of these expansive understandings:

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.

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. 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

---

164 Becker, supra note 161.
166 Perhaps because of the difficulty of defining the concept, many researchers (including Becker) refuse to define ‘human capital’, and instead focus on how human capital is created (i.e., the sources of human capital). The exception to this trend is most robust in governmental working groups and non-governmental organizations. Thus, of the three definitions used here, one is from the OECD, and the other is from a U.S. governmental working group with the following participating agencies - National Aeronautics and Space Administration, the Department of the Interior and the Department of Commerce. See Sustainable Development Working Group, Sustainable Development Indicators, Working Draft Framework, Version 2, June 4, 1996 available at http://www.hq.nasa.gov/iwgsdi/Human_Capital.html (last visited June 13, 2010): “Human Capital: The time, personal skills, capabilities, experiences, and knowledge of the individual. Human capital includes the entire life experience of the individual. Human capital does not include material or physical assets, which would fall under human-made capital or natural endowments.” See also OECD, MEASURING WHAT PEOPLE KNOW (1996) at 22: “Human capital is defined here as the knowledge that individuals acquire during their life and use to produce goods services or ideas in market or non-market circumstances.” See also MARTIN HUSZ, HUMAN CAPITAL, ENDOGENOUS GROWTH, AND GOVERNMENT POLICY (P. Lang 1998) at 9: “By human capital we mean the time, experience, knowledge and abilities of an individual household or a generation, which can be used in the production process.”
167 See e.g. PAGE, DIFFERENCE, supra note 70.
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. 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 ability to innovate, because it influences an individual’s capacity to perceive and create valuable opportunities in the environment. Perception affects both our ability to find problems, and to solve them.

Differences in preference and perception across individuals can often be a barrier to efficiency, as they raise the cost of coordinating activity, whether that coordination occurs within an organization, or between bilateral contracting partners. 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

---

168 This observation is echoed in NICHOLAS STERN, JEAN-JACQUES DETHIER, AND F. HALSEY ROGERS, GROWTH AND EMPOWERMENT: MAKING DEVELOPMENT HAPPEN (MIT Press 2005) at 244-258.
169 On how multiple preferences can create conflict, see PAGE, DIFFERENCE, supra note 70 at 11, 239-254.
170 For a discussion of how a multiplicity of preferences helps to solve problems, see PAGE, DIFFERENCE, supra note 70 at 256 and following: “Preference diversity begets toolbox diversity. And therefore, preference diversity has a strong and indirect effect on problem solving.”
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).

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.171 Others have argued that we will get innovations faster if we
have broader property rights, because the incentive for speedy commercialization is increased
when we protect business models with servitudes and license restrictions.172 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. This
perspective has much in common with the observations of Merges and Nelson on the impact
that patent scope has on innovation: “The only way to find out what works and what does not
is to let a variety of minds try.”173 However, it differs from their treatment because it extends
to the case of tangibles, which was explicitly disavowed in their landmark work: “unlike
rights that somehow touch tangible property… the allocation of property rights between
technological pioneers and improvers is not a zero-sum game.”174

This statement is consistent with the traditional analysis, and rightly so. Tangibles do
indeed appear to present us with a zero-sum game: if I have an object, you do not - end of
story. However, once we divide possession from use, the allocation of rights once again has
the potential to block the generation of new uses and the other innovations that arise from
asset use, and this effect is even greater when a reproducible tangible, rather than a unique
asset (i.e., chattels rather than land), is involved. The unification of use and possession over a
single, unique chattel may seem to be relatively insignificant when it comes to fostering user

171 See e.g. Robert P. Merges & Richard R. Nelson, On the Complex Economics of Patent Scope, 90 COLUM. L.
172 See e.g. Richard Epstein, The Disintegration of Intellectual Property? A Classical Liberal Response to a
Premature Obituary, 62 STAN. L. REV. 455 (2010) at 522: “But within this framework, we want individuals to
claim all potential uses that fall within the outer boundaries of their patent or copyright, subject to the other
limitations appropriate to the field. The more that creators can claim within these boundaries, the quicker these
innovations will be successfully commercialized.”
173 Merges & Nelson, supra note 171 at 873.
174 Id. at 877.
innovation. But once we consider that each user interacts with an asset differently from every other user (because of their unique stock of human capital), we can see that multiple copies of non-unique chattels can generate the same innovation enhancing benefits as non-rival goods. In other words, there is an underuse problem with tangibles as well as with intangibles. This fact has been recognized in a limited way by anticommons theory, which posits that dividing property among different owners can pose a barrier to efficient development if those fragments must be combined in order to make productive use of them. But combining property fragments is only one aspect of the unification of use and possession: far more important is the generation of new uses, asset modifications and other improvements that arise out of asset use.

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 intellectual property right can block these variations, and thus poses the danger of underuse. But once we see that each user interacts differently with an asset than every other user, we can see that chattels can have a similar multiplicative effect, and so allowing use restrictions on chattels therefore poses a similar danger of underuse. Although the scale of the effect is admittedly greater in intellectual property than with chattels due to the non-rivalrous nature of intangibles, the underlying dynamic is the same: just as many copies of an intellectual property work enable many users to develop the intangible resource, many copies of a tangible product enable many users to develop the tangible resource. Of course, the non-rival nature of the intangible resource means that only a single copy of the work needs to be sold in order to potentially enable an unlimited number of users to develop the resource; whereas many copies of a tangible chattel must be sold in order to enable many users to develop the resource.

The central role of reproduction in intellectual property tends to obscure the underlying similarities between copies or embodiments of intellectual property and more traditional chattels. This is because use restrictions in intellectual property often relate to an act of reproduction, but they need not be. At other times use is restricted even if no act of

reproduction occurs, as when a movie is performed publicly, for example. The underuse problem in intellectual property is generally attributed to pricing above marginal cost, which limits access to the good by those who value it at some value between the marginal cost of its production and the ‘monopoly’ price made possible by the presence of an intellectual property right. This effect is undoubtedly important in creating the specter of underuse, but the separation of use and possession also gives rise to an underuse problem that goes far beyond issues of pricing and access, and enters squarely into the realm of asset reuse (as when I watch a copy of a movie more than once) and derivative work creation (as when I cut photos out of old National Geographic magazines to create a collage for my bedroom wall).

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. But if there are many embodiments sold - as there are with most chattels and most copies of intellectual property works - then because each user interacts differently with an asset or copy than every other user, it is not necessary for each, or even the majority, of users to 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). Even though the potential trial and error process that results from unifying all rights of use and possession in an asset will generate many innovative dead ends, overall it will produce sufficient innovative outcomes to be worthwhile.

By unifying rights of use and possession whenever possible, the law of property organizes property rights around bundles of human capital. It is this feature which gives property law (at least in its numerus clausus incarnation) its innovation-promoting qualities. 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 unification

---

177 See e.g. F.M. Scherer, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 444 (2d ed. 1980).
178 Whether a collage infringes the derivative work right in any particular case depends on whether the collage ‘recast[s], transform[s], or adapt[s]’ the underlying works, 17 U.S.C. § 101. For one example where it did, see Jarvis v. K2, 486 F.3d 526 (9th Cir. 2007).
of rights of use and possession results in property rights being organized around their rival, nontransferable component. 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).

Of course, every allocation of property rights both reduces some costs, and imposes new costs of its own. And so it is with the law’s preference for unifying use and possession: we lose the benefits of the human capital inputs and decision set of the (now former) owner of the use right (let’s call that person T, for transferor), but we gain the benefits of the human capital inputs and decision set of the possessor of the asset (let’s call that person U1). If possessors and non-possessors had the same ability to search out new uses of assets, then this substitution would be offsetting at best. However, once U1 is already a possessor (by virtue of the market transaction which transferred possession, but not full use, from T to U1), then U1 has advantages in searching out new uses of that asset. Thus, as a result of the market transaction, we have already lost the benefits of T’s possession, but we have created the potential benefits that accompany U1’s possession (which we possibly forgo if we allow use rights to remain vested in T).

The case for unifying use and possession is even stronger when we consider a non-unique chattel, because in these cases we lose the benefits of T’s human capital and decision set, but we gain the benefits of the human capital and decision sets of every individual who purchases a chattel - say U1 to Un. Thus, the loss of the human capital inputs and decision set of T is a cost of unifying use and possession in users, but where possession has already been transferred, this loss is more than offset by the gain of the human capital inputs and decision sets of all those who now possess a copy of the asset (U1 to Un).179

CONCLUSION

Clearly, users have advantages over nonusers in seeking out new uses of assets that can be revealed only through asset use. This means that resource values will be higher as a

---

179 We should also note that the fact that users may innovate does not itself prevent the transferor from innovating, particularly in the case of a non-unique chattel, as he can simply retain copies or make additional copies. So, manufacturer innovation and user innovation are not mutually exclusive. What is important is that users be able to innovate, whether or not manufacturers themselves innovate.
result of user innovation. Just as clearly, a user’s cost of searching out new uses of an asset will be lower if he owns all rights of use and possession in the asset. But this lowered cost does not in itself explain why it matters that the user, rather than the nonuser, owns the right to permit a new use. If a user wants permission to search out new uses, why won’t a user and the owner of the use rights simply make a deal that reallocates those rights to the user, if that deal 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 indifference thesis). \(^{180}\)

The indifference issue has been hotly debated by both legal scholars and economists, and commentators have thus far identified two principal sources of transaction costs which may cause the thesis to break down: valuation and detection. \(^{181}\) 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. \(^{182}\) In the intellectual property case, it may be simply difficult to estimate how much value a particular intellectual property input adds to the value of a finished product. Such difficulties can cause bargaining to break down. Detection is a quite different matter: it may be difficult for an owner of intellectual property, for example, to know that a user has created a derivative work, made an infringing copy, or otherwise violated the intellectual property rights of the owner.

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 property law’s preference for exclusion. Asset valuation has generally been considered to significantly hinder transactions only when specialized assets (such as intellectual property and marital assets) are involved; \(^{183}\) therefore, this literature does not account for the broad range of assets and transactions regulated by the *numerus

---


\(^{181}\) Id. at 2656-60.


The various deficiencies of the standard account will be remedied in Chapter 3, where we move beyond disputes about valuation to examine the complex behavioral changes that arise when rights of use and possession in all sorts of assets - tangible and intangible, moveable and immovable – are divided. The most important effect of these behavioral changes, it will be shown, is that they alter the psychological ownership of assets, and this has a far greater impact on willingness to deal than mere asset valuation.

Problems related to detection, which implicate the ‘notice controversy’, will also be discussed in Chapter 3. Traditional economic wisdom posits that governance can generate the same innovative outcomes as exclusion, as long as the details of governance rules are fully disclosed to all parties. In other words, 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 exclusionary framework. So we need to ask ourselves: to what extent does notice of governance rules alleviate the concerns that animate the *numerus clausus*, thereby making the doctrine unnecessary? It is to this question, which marks an important point of departure between the user innovation theory and previous theories of the *numerus clausus*, that we now turn.

---

184 Other scholars have questioned the likelihood of successful bargaining on more general grounds: *see* Robert Cooter, *The Cost of Coase*, 11 J. LEGAL STUD. 1, 18 (1982). Cooter refers to this pessimistic view of bargaining behavior as ‘Hobbes Theorem,’ but he does not provide a general theory or guide as to when strategic behavior will interfere will value-enhancing deals. Chapter 3 will provide a related set of examples, and an explanation for them, for which the ‘Hobbes Theorem’ will hold true (i.e., when rights of use in an asset are separated from the right of possession).
Chapter 3: Psychological Ownership and Barriers to Recontracting

Nobody but a blockhead ever bought except for use.¹

INTRODUCTION

In Chapter 2, we examined why individuals interact differently with assets and produce different innovative outcomes, depending on how rights of use and possession are divided. We also examined the emergent nature of innovation, which makes it impossible to allocate appropriate rights of use to appropriate users ex ante to the innovative act.\(^2\) In this Chapter, we examine the impact that separating rights of use from the right of possession has on transacting behavior. Divided rights significantly decrease the willingness to deal on the part of both the owner of the use right and the possessor of the asset: this means that parties will frequently refuse to engage in ex post rearrangements of rights. The impact that divided rights have on transacting behavior has important implications for the role that notice should play in delimiting the boundaries of the numerus clausus.

In the sections that follow, we will examine why different allocations of use and possession cause individuals to transact differently over assets, greatly decreasing parties’ willingness to deal in situations of divided rights. The reason why traditional theory generally ignores problems related to the motivation to deal\(^3\) (i.e., to rearrange rights of possession and use) is relatively straightforward. Although notice generally operates to unilaterally bind downstream parties to restrictions (because no actual agreement to restrictions is required), 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.

The assumption that individuals will bargain rationally is woven deep into the fiber of economics. Nevertheless, disputes about an individual’s motivation to deal have arised now

---

\(^2\) In Chapter 2, we saw that the emergent nature of innovation means that no one (including the innovators themselves) can predict who will innovate, when they will innovate or how they will innovate. This means that no one (including the innovators themselves) can accurately judge ex ante when they should and should not give up their rights to innovate in exchange for a cheaper asset price, or when they should agree to ‘grant back’ such innovations to the seller of the asset. See Chapter 2, Part 1, iii) The Impossibility of Ex Ante Contracting. Once innovation has occurred, the behavioral issues discussed in this chapter will tend to hinder ex post recontracting.

and then even among economists. This is best illustrated by the sharp series of exchanges between Ronald Coase and Paul Samuelson about parties’ willingness to deal, 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.” The debate about motivation to deal in situations of bilateral monopoly spilled over into the legal academy, and gave rise to a great deal of scholarly analysis of bargaining behavior. That debate will not be revisited here. Instead, we will focus on the two sides of ‘motivation to deal’ issues that are present in the search for new uses of assets: 1) the motivation of the possessor of the asset to ask for permission for the new use, and to disclose the new use (the motivation to ask for permission/disclose); and 2) the motivation of the owner of the use rights to grant permission for the new use (the motivation to give permission).

**Part 1**

**THE MOTIVATION TO ASK FOR PERMISSION/DISCLOSE**

The motivations to ask for permission for a new use, and to disclose a new use, are obviously closely related, as the latter inevitably accompanies the former. Nonetheless, we will discuss each of these separately, because the need to ask for permission not only impacts the likelihood of disclosure to the owner of the use right; it also impacts the likelihood of free disclosure to the public. This latter effect has a particularly pernicious affect on the cycle of innovation, and so it deserves separate treatment. We first discuss the motivation to ask for permission, and then move on to the motivation to disclose.

---


A) The Motivation to Ask for Permission

In this section, we examine the impact that psychological ownership has on the transacting behavior of those who use and possess assets. In subsection 1, we see that psychological ownership interferes with the motivation of possessors of assets to ask for permission to use those assets, because possessors feel that they already ‘own’ the right to that use, via their use and possession of the asset. In subsection 2, we see that the psychological ownership that flows from the use and possession of an asset creates a strong expectation that full use of an asset should accompany the possession of that asset. Finally, in subsection 3, we examine evidence that the well-known ‘endowment effect’\(^6\) takes hold, not just upon legal ownership, but upon physical possession as well. The presence of an endowment effect in these circumstances means that even when users are willing to ask for permission for a new use of an asset, disagreements about the valuation of the use right might still hinder the successful conclusion of a deal.

1) Psychological Ownership, Possession and Use

When divided rights or governance rules are created, legal ownership of a physical asset is separated from legal ownership of the use rights in that asset. However, legal ownership is not the only form of ownership, and in its impact on human behavior, it may be the least important form of ownership. More important to human behavior is psychological ownership, which is generally defined as a state where an individual feels as though an object (tangible or intangible) belongs to him or her: ‘It is MINE.’\(^7\) Psychological ownership can co-exist with legal ownership, and indeed, legal ownership can help a state of psychological ownership to emerge.\(^8\) Nonetheless, legal and psychological ownership are two separate things, and they can operate in conflict with one another – i.e., an individual can have

\(^8\) Legal ownership can give a person socially sanctioned control over an object, and this control can lead to feelings of psychological ownership. See Id. at 96.
feelings of psychological ownership towards objects that he does not legally own. Such a situation is particularly likely to arise when legal rights of use are separated from possession, because as we shall see below, the processes via which individuals develop feelings of psychological ownership operate primarily through possession and use. As a result, the possessor of an object is highly likely to develop feelings of psychological ownership for the entire object, even though he is not the legal owner of all use rights in that object.

There are three routes to psychological ownership: 1) controlling an object; 2) getting to intimately know an object; and 3) investing the self into an object. Each of these routes can operate in isolation, or in tandem, and they tend to be both iterative and reinforcing. So, for example, being in control of an object can enable us to get to know an object, and provides an opportunity for us to invest the self into the object. In turn, the more we know about an object, and the more we invest the self into an object, the greater our feelings of control over that object. Control is both a route to psychological ownership, and a result of feelings of psychological ownership: we feel ownership towards that which we control, and we feel we should control that which we own. Most notable about the routes to psychological ownership is the fact that they are all implicated by the phenomena of possession and use.

Let’s first consider the ability to control an object. Possession is the most basic prerequisite to controlling an object; whereas use provides an ideal means for us to exercise and express that control. Indeed, asset modification and innovation can itself be considered to be the ultimate exercise of control over an asset: “The idea that we make things a part of the self by creating or altering them appears to be a universal human belief.”

In addition to control, knowledge is also an important component of psychological ownership. The more familiar we are with an asset, the greater our feelings of psychological

---

9 Id. at 86-87.
10 Id. at 92-93. See also Russell W. Belk, Possessions and the Extended Self, 15 J. CONSUMER RES. 139, 150-151 (1988) [Belk].
11 Pierce et al, supra note 7 at 96.
12 Id.
13 Other scholars have noted that use can be considered to be an exercise of control, see e.g. Lita Furby, Possessions: Toward a Theory of Their Meaning and Function Throughout the Life Cycle, in LIFE SPAN DEVELOPMENT AND BEHAVIOR 297 (P. B. Baltes ed., 1978) [Furby, Possessions]. Others have also identified use as a primary route via which objects become associated with the self, see e.g. CHARLES ELLWOOD, CULTURAL EVOLUTION: A STUDY OF SOCIAL ORIGINS AND DEVELOPMENT 184 (1927): “the fundamental element in ownership seems to have been use. The thing which was habitually used by a certain individual or group of individuals became assimilated to that individual or group.” (Emphasis in original.)
14 Belk, supra note 10 at 144.
ownership towards that asset will be. In speaking of this process, Pierce et al observe that “[p]eople come to find themselves psychologically tied to things as a result of their active participation or association with those things” in part because this association leads to knowledge and familiarity. This knowledge and familiarity leads to a sense of ownership of and connection to the object, and the object becomes incorporated into our sense of self.

Most significant for the user innovation theory of the *numerus clausus* is the fact that the process of repeated use and modification undertaken by users is ideally suited to generating an intimate level of knowledge of an object. Indeed, many users (particularly those who innovate) literally become experts in both the assets and their fields of use.

In addition to knowledge, the process of repeated asset use and modification provides an ideal opportunity for the user to invest the self into the object. Investing the self into an object can be as intangible as the act of creation itself, as when an author writes a book or when a user modifies his bike to fit his personal needs. In such cases, the self can become so tied into the object that it becomes part of an individual’s self-identity: ‘I am a writer’, ‘I am a surfer’, etc. However, users who ‘create’ are not the only ones who define themselves with respect to the objects that they use – collectors, for example, often define themselves with respect to the objects that they collect, and film and literary buffs are notorious for incorporating into their self-identity the books that they read and the movies that they watch.

Investing the self into an object can also occur via more seemingly impersonal routes, such as making personal sacrifices for the maintenance and upkeep of an asset that belongs to an employer. Truckers, for example, often invest their own time and money into personalizing and maintaining their employer’s vehicles, but only if they are assigned particular trucks. The assignment of a particular truck facilitates the exercise of control over, the development of knowledge of, and the investing of the self into, the truck. And these lead to feelings of ownership, which in turn lead to more investments, and so on. Here

---

15 “[W]e tend to prefer our own possessions to others, even others of a similar kind, because ‘we know them better’” (citations removed) Pierce et al, *supra* note 7 at 93.

16 Id. at 92.

17 Belk, *supra* note 10 at 144.


19 “The most obvious and perhaps the most powerful means by which an individual invests himself or herself into an object is to create it.” Pierce et al, *supra* note 7 at 93.

20 Clearly, in all of these cases, coming to know the object intimately also reinforces feelings of psychological ownership.

21 Pierce et al, *supra* note 7 at 96.
again, we see that the process of developing feelings of ownership is iterative, and that legal ownership and psychological ownership need not occur together.

It is of course true that psychological ownership will not vest in all assets, nor in all contexts, to the same degree. Most importantly, the development of such feelings can be expected to be somewhat diminished in the purely commercial context. Where assets are purchased for resale, the possession of and control over those assets is legal more than experiential, and such assets are not really ‘used’ at all. This general relationship has been postulated elsewhere as well. For example, it has been surmised that the endowment effect will not vest in assets purchased for resale, but only in those purchased for use. In the context of individuals innovating via assets purchased for use, however, psychological ownership can be expected to be at its strongest. The possession and use of assets gives rise to feelings of psychological ownership of the entire asset, not just the physical asset minus specified use rights. This has important implications for divided rights, because an individual will not ask for permission to use what he feels he already owns. The ability to control an asset is, after all, the most important aspect of ownership, and indeed, is virtually synonymous with ownership itself: “the two major components of the concept of ownership are control and attachment.”

2) Individuals Expect Use to Travel with Possession

The psychological ownership that arises from asset use and possession has an important consequence for those who want to implement use restrictions or other governance rules in respect to assets. Simply put, convincing people that they have to ask permission to use, or to incrementally pay to use, what they already possess, is an incredibly hard sell. Of course, individuals do ask for permission to use an asset whenever they rent, lease or borrow assets. However, in all of these cases, the act of asking for permission for use comes before the possession and use of the object – i.e., before feelings of psychological ownership have an opportunity to develop. Once the object is possessed and used, the duty to return the asset terminates any feelings of psychological ownership that may have developed during the term

of use and possession. 24 And, as we shall see below, even when individuals rent or temporarily possess assets, they want full use of the asset during the term of possession, even though they know they have to return it (i.e., even though they know that they do not ‘own’ it). 25 Indeed, the conviction that use of an asset should accompany its possession is so deeply ingrained that individuals will strenuously resist any interference with it. Needless to say, this creates an unwillingness to ask for permission to use a possessed asset, one which obstructs the process of contracting around governance rules.

In fact, the expectation that use travels with possession is so strong that it vests even when possession is only temporary. 26 In addition, the expectation persists even when an individual is fully apprised of use restrictions and when an individual explicitly agrees to them at time of asset purchase. That is, the expectation is not a mere by-product of lack of notice, or incomplete understanding of use restrictions. 27 We will discuss each of these features in turn, after we first examine some general evidence for the existence of the expectation, which we will generally refer to as the ‘use/possession expectation.’

### i) Evidence of the Expectation

The expectation that full use of an asset should accompany its possession has been occasionally noted by the courts: “in general, people both expect that and should be able to use the goods they purchase without restraints from merchants and manufacturers.” 28 Indeed, the existence of the *numerus clausus* can itself be considered to be evidence that judges often share this expectation. Some scholars have criticized courts for protecting this expectation, 29 but if individuals’ expectations shape both their interactions with assets and their transacting behavior, then it is in fact worth recognizing.

But the courts are not the only place that we find confirmation that the use/possession expectation exists. Perhaps the most direct evidence that individuals think that they should be able to fully and freely use the assets they possess can be found in the vociferous

---

24 Indeed, this explains why the duty to return is so pivotal in the law’s determination of when governance rules should and should not be enforced. See Chapter 4, Part 2, 2) Possession and the Duty to Return.

25 See below, ii) The Expectation Vests Upon Temporary as Well as Permanent Possession.

26 Id.

27 *Infra, iii) The Expectation is Not the Result of Cognitive Difficulties or Incomplete Notice.*


controversies that erupt in response to attempts to place use restrictions on assets. The ‘first sale’ doctrine,\(^{30}\) which gives the owner of a copy of a copyrighted work the ability to resell that copy, has more than once provided an effective focal point for public outrage against overreaching by intellectual property owners.\(^{31}\) But a past format war provides us with a particularly pointed illustration of the incendiary nature of use restrictions on assets.

\[\text{a) Of DivX and DVDs}\]

Most readers will not have heard of DivX technology. However, in the late 90’s, DivX was a pay-per-view format for renting and buying movies.\(^{32}\) The technology was essentially the same optical storage format as the DVD, but it came with encoding that needed to be unlocked by a central server, which was accessed via a phone call made by a modem within the DivX player. DivX players could play regular DVDs (the ones that we are by now well familiar with), but DVD players could not play DivX discs. Furthermore, the DivX player had to be connected to a phone line, not just to enable the initial viewing of the discs, but also to periodically check the user’s viewing history. If the DivX player was disconnected from the phone line, the discs would be locked until such a connection was re-established.

The process of ‘renting’ a DivX disc was similar to that of renting a DVD, but it had important differences. A consumer would select a disc from the Circuit City outlet, pay a ‘rental’ fee, and bring the disc home with them. At home, the consumer would place the DivX disc in the player, which would phone the central sever to unlock the encryption on the disc, and this would allow the consumer to freely view the disc for a period of 48 hours. Thereafter, if the consumer wanted to watch the disc again, he had two choices: he could either pay another rental fee, or he could pay to upgrade the disc to ‘Silver’ – which gave him the right to an unlimited number of future viewings. Both of these were accomplished by

\[^{30}\text{The first sale doctrine is codified in the Copyright Act, 17 U.S.C. § 109(a): “the owner of a particular copy or phonorecord lawfully made under this title, or any person authorized by such owner, is entitled, without the authority of the copyright owner, to sell or otherwise dispose of the possession of that copy or phonorecord.” The first sale doctrine is discussed in more detail in Chapter 5, Part 1, 2) B) i) First Sale.}\]

\[^{31}\text{See e.g. EFF Deeplinks Blog, First Sale, Why it Matters, Why We’re Fighting for It, August 9, 2007, http://www.eff.org/deeplinks/2007/08/first-sale-why-it-matters-why-were-fighting-it (last visited May 14, 2010).}\]

\[^{32}\text{In the following discussion, the factual details of the DivX technology, and its history, are drawn from: JIM TAYLOR, EVERYTHING YOU EVER WANTED TO KNOW ABOUT DVD 73-76 (2004).}\]
inserting the now-locked disc into the player and selecting the desired option; the individual’s DivX credit card would automatically be charged the appropriate amount. No matter what option the consumer chose (a single rental, multiple rentals, or unlimited viewings), the actual disc remained with the consumer – it never had to be returned.

The DivX format was developed by Circuit City, in an attempt to move towards that Holy Grail of the movie industry: pay-per-view implemented via physical copies. The value to the consumer was supposed to lie in the convenience of never having to return rented copies, and never having to pay late fees. There was also the convenience of being able to rent a disc, watch the movie, and then upgrade the disc to ‘Silver’ if they liked it enough to ‘buy’ it. Consumers could also opt to periodically pay for further viewings at their convenience (equivalent to going back to a physical store and re-renting a previously rented movie). There was also the convenience of a ‘rent now watch later’ option: because the initial rental fee was paid at the Circuit City outlet, but the 48 hour viewing window did not open up until the disc was actually put in the player, consumers could put the disc on the shelf and wait until much later to view it.

All of this sounds very convenient, but there were several glaring drawbacks to the scheme. Some of them have nothing to do with the use/possession expectation at issue here. Of more interest to us are the drawbacks related to the use/possession expectation, which helped raise public ire so much that the format did not even have a chance to gain market share. These drawbacks raised the specter of never owning copies, and they were implemented via legal agreement and by the technology of the DivX player itself. First,

33 The most troublesome shortcoming of the system was the ‘going out of business’ problem: because the discs had to phone home every few months to remain accessible, a consumer’s collection would become unusable if the company went out of business. In addition, the discs could only be played on a specially-bought DivX player, which had to be located near a phone line. Neither regular DVD players nor home computers could play DivX discs, while both could be used to play DVDs. In addition, there were a number of problems that resulted from the low market penetration of DivX, which existed because it was introduced after DVD had already gained a foothold the marketplace: the selection of movie titles was lower than that available on DVD; DivX discs often came with few or no ‘extras’ (such as interviews, etc.); the discs were only available at Circuit City; only a small number of manufacturers made DivX players; and there was also some dispute as to whether the initial rental price of the discs was somewhat too high. Most of these problems would have been remedied over time if DivX had gained sufficient market share (though the phone line problem would have remained, at least until wireless technology became widely available). Id.
34 I do not mean to suggest that the format’s interference with the use/possession expectation was the only cause of its failure. Clearly, the logistical problems referred to above, supra note 33, were also pivotal in the format’s demise. See e.g. Brooke Crothers, DivX DVD Backers Call it Quits, CNET News, June 16, 1999, http://news.cnet.com/2100-1040-227194.htm. But consumer-led anti-DivX campaigns, which used to be located at such sites as http://www.fightdivx.com/index2.html, did much to speed the format to its doom.
consumers never actually ‘bought’ a DivX disc, even if they paid to have it Silvered. Instead, they merely paid for permission to personally view it an unlimited number of times. Second, all uses of the disc were tied to a specific consumer. This meant that consumers could not freely lend a disc to a friend, nor could they resell a disc, even if they had paid for unlimited future viewings. Of course, a consumer could physically transfer a disc to a friend, either permanently or temporarily. But if the disc was locked (that is, if a consumer had merely paid for its rental), as soon as the disc was placed in the friend’s player, the player would phone the central server, detect an unpaid-for attempt to view, and charge the friend’s account. If the disc had been Silvered, placing the disc into the player would not immediately initiate a call to the central server, but the new disc in the friend’s collection would be detected during the monthly automatic account checks, and the friend’s account would again be charged. Clearly, there would be no reason for anyone to purchase used DivX discs: if they did, they would simply have to pay for the disc twice.

Thus, in the DivX universe, neither ‘lending’ nor ‘buying’ had the meanings that we traditionally ascribe to them. Indeed, even the concept of ‘renting’ was altered, because consumers never had to return discs, even if they only paid to view them once. But the absence of the need to return the copy was a bug, not a feature, because consumer retention of the copy raised the expectation of its full use. Or, in other words, permanent possession of the copy gave rise to feelings of psychological ownership, no matter what legal arrangement the consumer had agreed to.\textsuperscript{35} Thus, we see that consumers are willing to have restrictions placed on their ability to use an asset (such as when they have to return a rented disc at the end of its rental term), but only if they have to return that asset (which of course makes continued reuse impossible). Although DivX’s competitors also had much to gain from the failure of the format, the anti-DivX campaign was largely driven by consumers’ fury at the loss of ‘ownership’ that the DivX format implied: “The idea of a system designed for metered disc viewing enraged many film buffs and home theater hobbyists, and a virulent anti-DivX campaign erupted on the Internet.”\textsuperscript{36} Thus, the expectation that full and free use of an asset accompany its permanent possession is a powerful one.

\textsuperscript{35} In the case of ‘rental’, it was the right to view it for 48 hours; in the case of a Silvered disc, it was permission to view the disc an unlimited number of times.

\textsuperscript{36} R.J. Dunnill, \textit{A Short History of the DivX Home Video System}, DivX Owners Association, \url{http://www.thedoa.com/Pages/DoaDivxHistory.html}
b) Of DVDs and Droit de Suite

The impact of the use/possession expectation can be seen in the business as well as the consumer context. One example of this impact can be found in the history of resistance that DVD retailers display to movie studio demands for revenue sharing.\footnote{This resistance also shaped the history of the VCR and the VHS format. See the discussion infra Chapter 5, Part 1, 2) B) iii) The ‘Mere’ Right to Rent/Resell Fosters User Innovation.} In general, the history of the home video industry has shown that revenue sharing agreements tend to be successful when something is given to the rental business \textit{in exchange} for the shared revenue - such as privileged access to exclusive releases, or promised delivery of a certain volume of discs.\footnote{See e.g. Revenue Sharing Agreement between Blockbuster Videos Inc. and Columbia Tristar Home Video, Inc., August 25, 1998 \url{http://contracts.onecle.com/blockbuster/columbia_rsa_1998_08_25.shtml}.} Studio demands for revenue sharing are generally met with resistance, however, when it is demanded as part of a naked leasing scheme (i.e., when a studio demands a share of the revenue from DVD rental because the discs are ‘licensed rather than sold’).\footnote{The latest example of this type of dispute surrounds coin operated DVD dispensing machines. The low rental prices that these machines offer for new releases are seen as a threat to the studios’ ability to recoup costs via rental sharing agreements with large video rental outlets (whose rental prices for new releases are much higher). This new business model has caused some of the studios to demand that these vendors charge higher prices for the rental of new releases, and that they share this revenue with the studios. These demands have been backed up by refusals to sell new studio releases to the companies that operate these coin operated machines. One of the main players in this fledging industry, Redbox, has responded unfavorably to these demands. They have sued Universal Studios, seeking a declaratory judgment from the court that their activity is lawful under the first sale doctrine, and further asserting that Universal’s refusal to sell its discs to the company violates antitrust laws. The federal circuit has allowed Redbox’s suit to proceed. See Redbox Automated Retail LLC v. Universal City Studios, 2009 U.S. Dist. LEXIS 72700 (Fed. Cir. 2009).} This differential reaction shows that possessors of DVDs have a sense of ownership towards them: rental businesses are only willing to share rental revenue in exchange for some additional benefit (other than the mere right to rent the DVDs), because they feel that they already ‘own’ the right to rent the DVDs by virtue of their ownership of the discs.

It is instructive to contrast the resistance to revenue sharing in the sales of VHS cassettes and DVDs with the case of another resale royalty: the droit de suite. The droit de suite provides artists who create works of visual art with a share in the proceeds when their original artwork is resold.\footnote{See e.g. J.D. Stanford, \textit{Economic Analysis of the Droit de Suite – The Artist’s Resale Royalty}, 42 \textit{Aust. Econ. Papers} 386, 386 (2003).} The rationale behind such resale royalties is that early in their careers, artists must sell their works at low prices; once they have built up a reputation,
however, these early works may increase greatly in value.\textsuperscript{41} Because this increase in value is attributable to the artist’s continued efforts to create a body of work, and because artists who make a living from selling original works do not have the ongoing source of revenue that comes from selling multiple reproductions, it seems fair to allow such artists to share in the increased value of their early works.

The desirability of a droit de suite has been long debated in U.S. copyright law, but so far such a provision has not been imported into the U.S. Copyright Act.\textsuperscript{42} Nonetheless, both European countries and even some states have resale royalty laws,\textsuperscript{43} and once implemented, they do not appear to generate resistance to revenue sharing on the part of resellers (which has so frustrated the movie studios). At first this difference seems to cast doubt on the need to give purchasers free and full use of assets - but it actually has a relatively simple explanation. Most resale royalties do not apply to private sales; instead they only apply to works of art that are sold at auction.\textsuperscript{44} When works are sold at auction, the auction house, even without a resale royalty, gets a significant portion of the proceeds: this is the payment for the various services provided by the auction house. The resale royalty is thus seamlessly and painlessly folded into the already substantial fees that are charged for these original artwork transactions – i.e., the resale royalty in respect of an asset is hidden behind the provision of a related service.

The phenomenon of ‘hiding’ resale royalties can also be seen in the market for intangibles: when revenue sharing or incremental payments can be disguised by being folded into the provision of a service, the acceptability of ongoing payment schemes is relatively high. From the point of view of psychological ownership, the success of this somewhat surreptitious strategy makes sense: feelings of psychological ownership do not develop with respect to ongoing services, and so they do not interfere with these types of transactions. Similarly, when an asset must be physically returned after a certain period of time, such as when we rent a DVD, no feelings of psychological ownership attach to that asset. However, when ongoing payment schemes are implemented directly, via a demand for ongoing

\textsuperscript{41} Id. at 386-7.
\textsuperscript{43} Id. at 200-201
payments for assets that have been permanently transferred, such schemes are usually met with strenuous resistance. So, for example, the deactivation of a copy of a software program after a certain period of time, or upon contractual breach, has met with resistance.\(^{45}\) On the other hand, ongoing payment schemes are successfully implemented when no copy is transferred at all, such as when an online game provider sells access to an online game world.\(^{46}\) The fact that metered use pricing works well with the provision of services, rather than with the transfer of copies, provides us with more evidence of the power of possession, and of the practical impact that psychological ownership has on the development of successful business models. Just like the VHS and DVD rental entrepreneurs, consumers feel that they own a copy once it has been permanently transferred to them, and they do not want that ownership interfered with by copy deactivation or by ongoing payment obligations.

\[\text{ii) The Expectation Vests Upon Temporary as Well as Permanent Possession}\]

We have seen the impact that the use/possession expectation has on both consumer behavior and business decisions when assets are permanently transferred. However, the use/possession expectation is so strong that it vests, not just upon permanent transfer, but upon temporary transfer as well. This should perhaps not be surprising, because the presence of a traditional endowment effect was found in experimental subjects after an asset was possessed for only 15-30 minutes, and even though participants were informed that they did not own the asset.\(^{47}\) The DivX controversy clearly provides evidence that individuals expect use to travel with possession when they \textit{permanently} possess assets, but if we contrast the DivX system with the regular practice of movie rental, we can see that the expectation kicks in even when asset possession is \textit{temporary}. That is, even when the DVD must be physically returned to the owner, consumers expect to be able to view the movie as many times as they want during the period in which they physically possess it. This is so even if in practice, they


\(^{46}\) See \textit{e.g.} World of Warcraft FAQ, \url{http://www.worldofwarcraft.com/info/faq/general.html} (last visited June 15, 2010).

only view it once (or not at all). As a result, consumers would rather have the inconvenience of returning a DVD that they may only watch once, than have the convenience of never returning a DVD at all. The desire for the mere option of repeated use of a rented asset shows that individuals will act like owners (at least during the term of asset possession) even if they know they have to return an asset. 48

**iii) The Expectation is Not the Result of Cognitive Difficulties or Incomplete Notice**

Before we conclude our discussion of the use/possession expectation and the effect that it has on human behavior, we need to differentiate it from problems that arise from a lack of consent or notice. Consent/notice problems can be organized into three rough categories: those that arise from technological limitations in the mechanics of providing notice; those that arise from cognitive insufficiencies on the part of consumers; and those that are created by a mismatch between the parameters of the asset and the degree of the use restriction. The first two of these problems have long been recognized in the literature on standard form and adhesion contracts: “the adhering party is in practice unlikely to have read the standard terms before signing the document and is unlikely to have understood them if he has read them.” 49 In the case of limitations in the mechanics of providing notice, a consumer will not even be aware of restrictions, and this makes it impossible for him to agree to them. In the case of cognitive difficulties, a consumer may have some sort of actual notice, but the restrictions may be so complex that he is unable to interpret them.

A more recent theory proposes that, in some cases, use restrictions lack salience. 50 In these cases, buyers are aware of use restrictions, but they are unable to sufficiently assess their impact at the time of purchase because they form a relatively minor part of a much larger property bundle. 51 This may happen in the case of a real property transaction, for example. A buyer focused on the large lot of land and the new house he is purchasing might

---

48 To the extent of repeated use, but not to the extent of transfer or destruction.
51 Id. at 1228-1229.
not pay much attention to an easement in the deed; but what might appear insignificant at the
time of purchase might later turn out to be inconvenient, even irksome.\textsuperscript{52}

These rationales for setting aside use restrictions doubtless have an important role to
play in a variety of transactions; however, the expectation that full use travel with possession
(and the behavioral responses that accompany it) has nothing to do with lack of notice,
cognitive inability, or salience. In others words, even when individuals actively agree to use
restrictions with full awareness and comprehension of their legal obligations, once they
possess and use an asset, they will tend to develop feelings of psychological ownership.
Indeed, this will happen even when they make a positive choice between a higher-priced
unrestricted good, and a lower-priced restricted good. An example of this dynamic can be
found in the case of non-refillable laser printer cartridges, like the ‘Prebate’ cartridges sold
by Lexmark.\textsuperscript{53} Consumers are offered the choice between a lower-priced non-refillable
cartridge and a higher-priced refillable cartridge, but they routinely insist on purchasing and
then refilling the lower-priced cartridges. Consumers are well aware when they make the
choice to buy the cheaper, well-marked cartridge that they are agreeing not to refill the
cartridge, but they insist on doing so anyway. Now, it is true that this overwhelming
willingness to ‘cheat on the deal’ might simply be the result of opportunistic behavior.
However, even if some consumers intend to abide by the restrictions on use at the time of
purchase, subsequent use and possession of the cartridge is highly likely to give rise to
feelings of psychological ownership which displace that intention. In other words, mere
notice of use restrictions is unlikely to be sufficient to prevent the development of feelings of
psychological ownership that flow from use and possession.\textsuperscript{54}

3) Psychological Ownership and the Endowment Effect

The ‘endowment effect’ refers to the phenomenon that occurs when the price that we
are willing to pay (willingness to pay, or WTP) for an asset that we do not own is less than

\textsuperscript{52} See e.g. Mark Kelman, A Guide To Critical Legal Studies 108-09 (1987).
\textsuperscript{53} For this and other examples, see generally the discussion in: Christopher Soghoian, Caveat Venditor:
Technologically Protected Subsidized Goods and the Customers Who Hack Them, 6 NW. J. TECH. & INTELL.
PROP. 46 (2007).
\textsuperscript{54} We return to the discussion of what is sufficient to prevent or counterbalance such feelings in Chapter 4, Part 2, 3) c) ii) Psychological Ownership, Notice and the Bargaining Process.
the price that we are willing to accept for an asset that we do own (willingness to accept, or WTA).\textsuperscript{55} According to traditional economic theory, there should be no difference between these two amounts: if you are willing to pay up to 10 dollars for an item, you should be willing to sell it for as little as 10 dollars. But the presence of an endowment effect means that although you are willing to pay up to 10 dollars for an item, once you own it, you are no longer willing to accept a payment of as little as 10 dollars to sell the item, but instead require more.

The endowment effect is distinct from the type of ‘personhood’ interest\textsuperscript{56} which tends to attach only to irreplaceable objects, such as wedding rings, to which we develop a specialized emotional attachment. In contrast, the endowment effect has been shown to attach to even simple, fungible assets (like coffee mugs), and so it is potentially relevant even where garden variety chattels and other readily replaceable assets are involved. In addition, personhood effects tend to move assets from the market to the non-market realm (where we would not sell them for any price), rather than merely increasing the price that we demand for them.

The presence of the endowment effect is important because it 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 then 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, 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.

Traditionally, the endowment effect has been analyzed in regards to legal ownership, but research has shown that the endowment effect arises, not just from legal ownership, but from \textit{mere possession} as well.\textsuperscript{57} A recent experiment measured this effect by giving one

\textsuperscript{55} Richard H. Thaler, \textit{Toward a Positive Theory of Consumer Choice}, 1 J. Econ. Behavior & Org. 39 (1980). The existence of endowment effect has been hotly debated. That debate will not be examined here. \\
\textsuperscript{56} Margaret Jane Radin, \textit{Property and Personhood}, 34 STAN. L. REV. 957, 959 (1982) [Radin, Personhood]. \\
\textsuperscript{57} Reb & Connolly, \textit{supra} note 47. It should also be noted that in their experiment, Reb and Connolly did \textit{not} find an endowment effect when participants were given legal ownership, but not possession, of chocolate bars and mugs. Although this implies that the endowment effect only takes hold upon possession and not with legal
group of participants possession of a chocolate bar for 15-30 minutes while they took part in an unrelated study. They were told that it would be used in a later study, and to simply hold on to it until then. After the unrelated study was completed, they were told that they could acquire the chocolate bar or receive a sum of money. Another group of participants, who also took part in the unrelated study, were not given possession of the chocolate bar, but were shown it briefly after the unrelated study was completed, and were told that they owned the bar and could exchange it for money if they wished. Participants then answered a series of questions about their preference between keeping the chocolate bar and exchanging it for various amounts of money (or, in the case of the possession/no ownership participants, between acquiring the chocolate bar and receiving various amounts of money). They were also asked a series of questions about how much they felt they owned the chocolate bar (whether they were in the group that actually owned it or not). The results clearly showed that “the sense of endowment that leads to higher monetary valuations results from the feelings of ownership induced by possessing an object, rather than legal ownership as such.”

In the case of possession, the endowment effect arises because possession itself gives rise to subjective feelings of ownership, and these feelings change how possessors behave. In fact, possession may be even more important than legal ownership in determining the existence and the degree of endowment effect that attaches to a property right.

The fact that mere possession of an asset can give rise to feelings of ownership and an endowment effect means that, despite their lack of legal ownership, the possessors of assets sometimes behave as if they own those assets. Of course, when it comes to owning a tangible object, but not all rights of use of that object, the endowment effect, at least in its traditional incarnation, doesn’t really seem to apply. In these situations, it seems somewhat inappropriate to speak of valuing an asset more highly, because the possessor of the tangible object is put in the position of valuing an asset (the use right) that he literally does not own. It

ownership, it is likely that experimental design caused this discrepancy. In this experiment, the legal owners never had possession, and the possessors never had legal ownership. However, in most cases, use rights are retained by sellers who are previous possessors of the asset. Furthermore, in the case of very valuable assets, such as intellectual property works and stock options, an endowment effect takes hold even thought there is no tangible asset to possess. See also Amitai Etzioni, *The Socio-Economics of Property*, 6 J. SOCIAL BEHAVIOR & PERSONALITY 465 (1991); and Furby, Possessions, *supra* note 13.

58 Reb & Connolly, *supra* note 47 at 110.
59 Psychological feelings of ownership are generated in three ways: controlling an asset, becoming familiar with an asset, and investing the self in an asset. See Pierce et al, *supra* note 7 at 92-93.
60 Reb & Connolly, *supra* note 47 at 108.
is therefore nonsensical to speak of a possessor demanding a higher price for a use right which in actuality belongs to another. Nonetheless, possession of an asset does give rise to feelings of psychological ownership, which hinder the process of contracting around governance rules because they make possessors act like owners.

As a result of the endowment effect of possession, the already existing reluctance to ask for permission to use a possessed asset is further exacerbated. In other words, an individual who permanently possesses an object not only feels that he already ‘owns’ the use (because he feels that he ‘owns’ the possession), but he also feels that he should be compensated even more for the possession (and by extension, the use) than he paid for the possession (even though he does not legally own all rights of use in that possession). Thus, dividing use from possession creates a situation in which the possessor of an asset is legally required to ask for permission to do (and to pay for) something which he feels he already owns, and for which his WTA is higher than his WTP (even though his WTA arises in regards to something he does not legally own). And even when an individual asks for permission in regards to an asset that he already possesses, the endowment effect of possession can interfere with the successful conclusion of a deal, because it can generate disputes over valuation of the use rights.

B) The Motivation to Disclose

It may seem redundant to discuss the motivation to disclose a new use after we have already discussed the motivation to ask for permission for that new use. However, in this section we are interested in the motivation to disclose the new use, not to the owner of the use right, but to the public at large. In fact, we are concerned not just with disclosure, but with the willingness to freely disclose: to give the world at large access to the innovation, and to allow others to reproduce, imitate and build upon that innovation if they so choose. Public disclosure is important because the existence of a legal use restriction may not itself stop user innovation. A user may simply refuse to ask for permission to violate a use restriction, and so the innovation will occur despite the existence of that restriction. This is particularly the case with chattels, as it is generally difficult for the owners of use rights to monitor asset use and detect violations.
However, the practical impact of a use restriction can show itself when it comes time for a user to disclose a new use to the public. That disclosure can be active, as when a user posts information on a website; or it can be passive, as when a surfer uses his modified board at the local beach and fellow surfers take notice. Either way, if a user is worried about getting sued for violating a use restriction, then this may encourage him to keep the use private. This interrupts the cycle of free revealing, add-on innovation, and commercial adoption that is critical to the innovative impact of asset use. Free revealing is not akin to throwing something away: it is the bottleneck in an uncoordinated, but extremely valuable process of incremental innovation.

Recently, a substantial literature has built up around the phenomenon of free revealing or free disclosure, most of it in response to the prominence of the ‘open source’ development of computer software. Although the scholarly awareness of open source has increased along with the growth of networked environments that both facilitate information sharing and make it visible to academics, free disclosure has always played an important role in the innovative process. Sometimes, it took the form of defensive publishing: when a company did not plan to patent or develop an innovation, it would publish in a scientific or trade journal in order to prevent a rival from getting a patent. At other times, free disclosure took the form of norms of reciprocal information sharing in an industry, such as the regularized practice of free disclosure that occurred in the iron industry in England in the nineteenth century, and in the early development of mine pumping engines. And sometimes, it was simply impossible to keep information secret.

---

61 See Von Hippel, supra note 18 at 77.
63 Id.
64 See e.g. Ilkka Tuomi, Networks of Innovation: Change and Meaning in the Age of the Internet (Clarendon 2002); and Open Innovation: Researching a New Paradigm (Henry William Chesbrough, Wim Vanhaverbeke, Joel West, eds., 2006).
But the highly visible phenomena of the Internet and open source communities have caused an upsurge of interest in free disclosure. This literature has identified many additional reasons why users and other innovators might choose to freely disclose their innovations, rather than seeking to profit from that disclosure.\textsuperscript{69} For example, there might be reputational advantages to disclosing prowess to a specialized community.\textsuperscript{70} Furthermore, if someone else might eventually hit upon the same innovation, there is an advantage to freely disclosing to your peers, in order to get the credit for ‘being first.’\textsuperscript{71} Just as there are other ways for manufacturers to gain advantages in the marketplace, even without intellectual property rights (such as lead time advantage, tacit knowledge, and production know-how), there may be other ways for users to extract value from disclosure, even without property rights and monetary payments.

The demonstrated willingness of users to freely disclose their innovations has important implications for innovation policy, for it implies that the disclosure problem in intangibles is overstated. Indeed, the phenomenon of user innovation itself implies that the production problem in intangibles, which provides the rationale for granting intellectual property rights, is overstated. If a significant amount of innovation occurs as a by-product of asset use, there is little need for an intellectual property right to provide incentive for its production. And if disclosure is often free, then the disclosure argument for intellectual property - which posits that even if intellectual property does not actually provide an incentive for production, it provides an incentive for disclosure because it solves Arrow’s information paradox\textsuperscript{72} – may also be somewhat overstated. The prevalence of user innovation implies that we may not have to provide an incentive for the production and disclosure that arises out of the organic process of asset use, but simply allow it. Allow it to be freely produced, and allow it to be disclosed, freely or otherwise, at the discretion of the

\textsuperscript{69} Paul Muller & Julien Penin, \textit{Why Do Firms Disclose Knowledge and How Does it Matter?}, in \textit{INNOVATION, INDUSTRIAL DYNAMICS AND STRUCTURAL TRANSFORMATION} 149, 150 (Uwe Cantner & Franco Malerba, eds., 2006): “Since such behaviors of open knowledge disclosure are costly and do not give rise to direct rewards, they have often been misunderstood by mainstream economics.”


\textsuperscript{71} See \textit{VON HIPPEL}, \textit{supra} note 18 at 77-91.

\textsuperscript{72} Arrow’s information paradox arises when the asset over which parties must transact is largely or wholly comprised of information. In order for the parties to strike a deal, the information must be revealed from the seller to the potential buyer; however, once the information is revealed, the potential buyer has the information, and so there is no longer any reason for him to strike a deal. Arrow, \textit{supra} note 3. If parties are willing to freely reveal information, however, Arrow’s information paradox disappears.
user. The best way to do this is to unify rights of use and possession in assets, because 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.

This leads us to the third assertion about user innovation:

3. 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.

Now that we have discussed the various phenomena that can interfere with the motivation of the user to ask for permission from the owner, we will examine the motivation of the owner to grant use rights to the user.

Part 2 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 rational actor model, it is far more common than is often acknowledged.73 We will discuss four phenomena that have a negative impact on the motivation of economic actors to give permission for a new use, even when presented with an efficient transaction:74 1) psychological ownership and the ability to refuse permission; 2) the endowment effect of legal ownership; 3) noneconomic motivations; and 4) 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 (shortsightedness/incumbent myopia). The first phenomenon is especially relevant when rights of use are divided from the right of possession, and is particularly operative when a use right is left in the hands of a former possessor of an asset. The other three are general to all transactions, but are nonetheless worth discussing because they contribute to the overall reluctance of use right owners to grant permissions for new

73 Of course, there will be times when the refusal to give ex post permission is rational. These cases are of little concern to the numerus clausus (though they may be of great concern to intellectual property), because they involve disputes over value rather than disputes over ownership. See infra Chapter 3, Part 4, Value Versus Ownership.

74 There may, of course, be more phenomena. But these are taken to be those that are of particular importance.
uses. Thus, only with a consideration of all of these phenomena does a complete picture of refusals to deal emerge.

1) Psychological Ownership and the Ability to Refuse Permission

Even after psychological feelings of ownership have developed with respect to an asset, they are not necessarily permanent. If the routes to psychological ownership are blocked or removed, then feelings of psychological ownership can also diminish over time, and even cease altogether.\(^{75}\) The process via which feelings of psychological ownership diminish over time (this can be either a short or long period) is referred to as the ‘decoupling process.’\(^{76}\) At times, we even engage in formal rituals to hasten this ‘decoupling process’ between the self and the object of ownership.\(^{77}\) In the case of an asset that is ‘sold’ on the market, the physical transfer alone, because it involves the loss of control of the asset and the cessation of the investing of the self, is usually sufficient to decouple the object from the sense of self (thereby eliminating feelings of psychological ownership). Indeed, even the decision to sell an asset is likely to begin the decoupling process, because the decision to sell puts some distance between the asset and the seller’s sense of self. This is why psychological ownership operates differently in the case of a use-restricted asset that is not for sale, and a used asset that is for sale. That is, feelings of psychological ownership may increase the asking price for a used asset, thereby creating a dispute over value, but they do not create a dispute over ownership. In contrast, in the case of a use-restricted asset, feelings of psychological ownership create a dispute over ownership.\(^{78}\)

These disputes over ownership have already been seen from the perspective of the individual who possesses an asset but who must ask for permission to use it. However, these disputes over ownership must also be viewed from the perspective of the individual who must grant permission for such use. The reason for this is that the transfer of possession and the retention of use rights have conflicting effects on feelings of psychological ownership.

---

\(^{75}\) Pierce et al, supra note 7 at 97.
\(^{76}\) Id.
\(^{77}\) Pierce et al list some of these rituals: “estrangement, divorce, devaluation, hostility, depersonalization.” Id.
\(^{78}\) The distinction between disputes over ownership and disputes over value is discussed in more detail later in this chapter. See Chapter 3, Part 4, Value Versus Ownership.
The transfer of possession to the user creates feelings of ownership in the user, and eliminates feelings of ownership in the seller: these effects reinforce each other and do not create conflicts over ownership. However, when an asset is physically transferred but use rights are maintained, the transfer of possession gives rise to feelings of ownership in the user, but the retention of use rights also maintains feelings of ownership in the seller. These effects antagonize one another and create conflicts over ownership. In other words, the retention of use rights in the asset hinders the decoupling process that would have occurred if the seller had parted with full ownership of the asset.

There are two reasons why the decoupling process is hindered when the seller retains use rights. First, legal ownership itself gives rise to feelings of psychological ownership. Second, the ability to deny permission for use itself gives rise to feelings of psychological ownership, because the ability to deny individuals permission to use an asset is one of the key incidents of psychological ownership. For example, when children are the first to hear a nursery rhyme or to use an object, they feel that this gives them ownership of the rhyme or object, and consequently, they feel that other children cannot hear the rhyme or use the object without their permission. The ability to refuse permission to use an asset, because it is such a fundamental aspect of ownership, is likely to preserve the feelings of psychological ownership with respect to a transferred asset.

Research has also shown that ownership is not simply about control over objects: it is also about control over others. Thus, our understanding of ownership begins with the spatial connection between an object and a person (i.e., ownership is associated with possession and use of an object), but it progresses from this to direct control of an object, to the control of use of an object by others, and finally to direct control of others. Thus, ownership is not just about control over assets: it is also about control of others via the control of assets. The ability to control others via the refusal of permission to use assets is therefore a key indicium of ownership.

Whether ownership is perceived as the ability to control objects, or others, or both objects and others, the most effective way for a use right owner to exercise control over the
asset may be to deny permission for the new use altogether, because the use right owner cannot herself use the asset.\footnote{No studies have been done on whether a monetary payment is sufficient substitute for an exercise of control.} Thus, feelings of psychological ownership may lead to the denial of permission even in the face of value enhancing deals.

2) The Endowment Effect of Legal Ownership

We have already seen that the endowment effect can interfere with the motivation to ask for permission for a new use: possession of a physical asset can give rise to feelings of psychological ownership, and these feelings of ownership can give rise to the endowment effect.\footnote{Reb & Connolly, supra note 47.} But the endowment effect can also interfere with the granting of 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. Indeed, the supply side of permissions presents us with the more traditional setting of the endowment effect - which involves legal ownership, rather than mere possession.\footnote{Id. at 108.} In the case of permissions, the object of legal ownership is a use right, and so the object of ownership is intangible in nature. Although some studies have indicated that the endowment effect is 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\footnote{Thaler, supra note 55.}), other studies have also observed the effect in cases of ‘inherently’ intangible assets, like stock options and intellectual property.\footnote{Cynthia Devers, Robert Wiseman & Michael Holmes, The Effects of Endowment and Loss Aversion in Managerial Stock Option Valuation, 50 ACADEMY OF MGMT. J. 191 (2008); Christopher Buccafusco and Christopher Sprigman, Valuing Intellectual Property: An Experiment (March 11, 2010). Chicago-Kent Intellectual Property, Science & Technology Research Paper No. 10-029; Virginia Law and Economics Research Paper No. 2010-04. Available at SSRN: http://ssrn.com/abstract=1568962.} Thus, we would also expect an endowment effect to take hold in other intangible legal rights, such as a right of use.

It may be objected that, in the case of intellectual property at least, what appears to be an endowment effect is actually a manifestation of the personhood interest that arises from the act of creation.\footnote{See e.g. Roberta R. Kwall, Inspiration and Innovation: The Intrinsic Dimension of the Artistic Soul, 81 NOTRE DAME L. REV. 1945, 1971-73 (2006).} Although we would undoubtedly expect the endowment effect in intellectual property to be interwoven with the complex personhood effects that accompany
the act of creation or innovation, there is ample evidence that corporate assignees of intellectual property can be far more intransigent in their dealings over intellectual property than authors or inventors.88 This is not to say, of course, that authors and creators are always reasonable at the negotiating table, but when a mere owner of intellectual property displays such obstinacy, we cannot attribute it to the emotional attachment that comes from creation. In addition, we have already seen that possessions that are purchased89 or found90 (whether tangible or intangible), are just as central to personhood as ‘creations’ (whether tangible or intangible).91 Indeed, some thinkers push the equivalency much further than this: “Sartre feels that buying an object is merely another form of creating the object.”92

3) Noneconomic Motivations

If sellers of assets were always driven by economic motivations, changes in price would motivate sellers to impose only those restrictions which are desirable to prospective buyers. If a seller 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 by lifting the restrictions to make the asset more marketable. If the market weeds out restrictions that are of

88 Institutional intellectual property owners regularly walk away from profitable licensing deals, for no apparent economic or creative reason. For example, when Jon Else, a documentary filmmaker, wanted to license a short clip of the Simpson’s cartoon which was playing on a television set in the background of a scene, he got the go-ahead (gratis) from both the cartoon’s creator and the production company. However, the institutional owner of the copyright in the work, Fox, demanded 10,000 dollars for the 4.5 second clip. This demand was made with full knowledge of the fact that the film was a documentary about the San Francisco Opera which had nothing to do with the Simpson’s. Thus, there was no dispute about the ‘value’ that the clip added to the documentary. There was also no concern that use of the clip would cast a negative light on the show. After checking and rechecking this demand with top executives, Else removed the clip from his documentary.

89 See our discussion of the DivX discs, supra notes 32 to 38 and accompanying text; also see Belk, supra note 9 at 160 (commonly purchased items).

90 Like the chocolate bars in Reb & Connolly, supra note 47.

91 In reality, there is likely some sort of ownership continuum which has objects that are incorporated into the sense of self at one end, and objects that we legally own but for which we do not have feelings of ownership at the other end. This accords with the observations in Pierce et al, supra note 7 at 87. This continuum has objects on it that are created/purchased/found, tangible/intangible, common/rare and fungible/nonfungible, but all of these types of objects can be located at any point on the continuum.

92 Belk, supra note 10 at 150.
little economic value, it is inappropriate and unnecessary for the law to interfere and second
guess the market.

Certainly, the market does much to weed out undesirable products, but asset
restrictions do not seem to be as responsive to consumer preferences. Asset restrictions often
become standardized across industries and across assets, and this makes them much less
responsive to price pressures.\(^3\) For now, we will consider the more discrete, bilateral
instance of this phenomenon: what happens if sellers, even if they have complete information
about consumer preferences, do not respond to price changes? That is to say, what if they are
driven by noneconomic motivations? 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). One well-
known situation in which this would often occur (but for legal intervention) is in the case of
parody. A parody is an artistic work which builds upon, and to some extent mimics, another
work for the purpose of mockery. Parody is a form of criticism, albeit a rather specialized
one. It is easy to see why owners of original works might not grant permission to parodists,
even at generous prices. Quite simply, we may be unwilling to allow others to make fun of us
(or our artistic children) for any price.\(^4\) Thus, if the law required parodists to get permission
to parody, then such works might appear rarely, if at all.\(^5\) The law of copyright circumvents
this problem by removing the right to authorize parodies from the bundle of rights legally
granted to the copyright owner.\(^6\) Normally, the creation of derivative works requires the
permission of the copyright owner of the original work, but in the case of parody (as well as

---
\(^3\) Florencia Marotta-Wurgler, *Competition and the Quality of Standard Form Contracts: The Case of Software

\(^4\) See e.g. Wendy Gordon, *Fair Use as Market Failure*, 82 COLUM. L. REV. 1600 (1982); and Robert P. Merges,
*Are You Making Fun of Me? Notes on Market Failure and the Parody Defense in Copyright*, 21 AM. INTELL.

\(^5\) Although some might say that this implies that the value to the parodist of producing the parody is less than
the value to the owner of the original work of not having a parody, copyright law generally presumes that the
social value of the parody outweighs the harm to the author of the parodied work. Indeed, it is more accurate to
say that copyright law, via the fair use doctrine, removes some forms of expression from the economic calculus
entirely, thereby expressing society’s noneconomic motivations or judgments.

in other types of comment and criticism) the permission of the owner of the copyright in the commented-upon work is not required.

Another common example in which noneconomic motivations pose a pricing problem occurs in the case of government expropriation of housing and land. Some commentators have suggested that resistance to expropriation is best understood in terms of mere undercompensation, rather than as a case in which money is an inadequate substitute for the property. Nonetheless, in many cases, there seems to be evidence that a personhood-type interest can, under the right conditions, vest in these types of assets, making individuals less responsive to price changes.97 Finally, land conservation, historical preservation and environmental protection easements are legal devices explicitly designed to accommodate noneconomic values.98

Noneconomic motivations are extremely problematic for the rational actor model, but mainstream economists are becoming more aware of the need to acknowledge the existence of noneconomic motivations: “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.”99 This recognition of the problem is heartening, but it may underestimate its prevalence. Evocation of the ‘zealot’ may be comforting because it implies that the phenomenon is relatively rare. After all, which of us, in the ordinary course of things, can afford the cost of unadulterated zeal? But what if zealotry were not a rare wrinkle in economic theory, but rather a common human characteristic? Perhaps there is a bit of the zealot in all of us, and perhaps it simply takes the right set of circumstances (e.g., type of asset) to bring it out.100

97 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. See e.g., Charles E. Cohen, Eminent Domain After Kelo v. City of New London: An Argument For Banning Economic Development Takings, 29 Harv. J.L. & Pub. Pol'y 491 (2006).
99 See e.g. James Buchanan & Yong Yoon, Symmetric Tragedies: Commons and Anti-Commons, 43 J. L. ECON. 1, 12 (2000).
100 Indeed, perhaps it only takes the right person to bring it out. One infamous example of such a refusal to license comes to us from the aviation industry: the Wright brothers’ refusal to license the best and most
4) Shortsightedness (Incumbent Myopia)

Although we saw in Chapter 2 that manufacturers often recognize the value of user innovations and begin to manufacture them, the relationship between innovating users and industry is not always a harmonious one. This is especially the case when user innovations appear to disrupt or threaten the established business models of incumbents. Consider Hollywood’s resistance to the rental of VHS cassettes. It may seem difficult to believe\(^1\), but the early reaction of the major studios was a mixture of hostility, fear and disbelief. One reason for this reaction was that they were so accustomed to leasing, rather than selling, movie prints to theatres that they simply could not see any other way of doing things. The thought of distributing copies of their movies, even on a downgraded format like VHS, seemed tantamount to losing complete control over their copyrighted works. They were accustomed to taking part in (i.e., receiving revenue from) each and every viewing of a film, from each and every pair of eyes, and they did not want to give that up. When RCA demonstrated an early version of non-rewindable magnetic tape to Disney executives: “The executives were horrified. They would ‘never’... permit their content to be distributed in this

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, but they did eventually license various entities. However, there was one individual whom Orville steadfastly refused to license - Glenn Curtiss: “Wright... announced that he will consider lenient royalty arrangements with anyone in the field except Curtiss.” (Emphasis in original.) Who was Glenn Curtiss? He was an inventor, 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, 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. 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 should not be underestimated. 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. See SETH SHULMAN, UNLOCKING THE SKY (Perennial 2003) (2002).

\(^1\) This is especially surprising when we consider that studio revenue from the sale of VHS cassettes (and later DVDs) to rental shops went on to eclipse box office receipts, and finance the blockbuster model of movie production. And it did this by supplementing, rather than cannibalizing, box office revenue. For a brief discussion of this history, see Chapter 5, Part 1, 2) B) iii) The ‘Mere’ Right to Rent/Resell Fosters User Innovation. For a complete history, see FREDERICK WASSER, VENI, VIDI, VIDEO (University of Texas Press 2001).
form. For the content, however clever the self-locking tape player was, was still insufficiently controlled. ‘How could they know,’ a Disney executive asked… ‘how many people are going to be sitting there watching’ a film? ‘What’s to stop someone else coming in and watching for free?’”102

The distribution of copies for home viewing also raised serious questions about the movie industry’s business model, which was based on bringing people into the theatres. Would the public still go out to the theatre if they could simply watch movies at home? The threat that home viewing would cannibalize, rather than complement, the theatre run likely explains why studios were even more hostile to renting cassettes to consumers, than to selling cassettes to them. Serious movies buffs would be likely to both see a movie on the big screen and add it to their high-priced collection; a casual viewer, on the other hand, might use rental as a substitute for a trip to the theatre. Losing control over copies also raised fears of piracy: the VCR’s attractiveness hinged on its ability to make copies of copyrighted works. Although it was marketed as a time-shifting device which consumers could use to record their favorite television shows for later viewing, it did not take much imagination to see that if VCRs could be set up in tandem, consumers could bootstrap a cheap rental into a permanent copy.

For all of these reasons, in the early days of home video, the major studios spent more time and money trying to hinder the spread of rental shops and the market penetration of VCR technology, than they did attempting to capitalize on this substantial new source of revenue. The first sale doctrine, which gives the owner of a copy of a copyrighted work the ability to resell and reuse that copy, allowed rental shops to rent videos without the permission of the studios.103 By unifying the rights of reuse and resale with the right of possession of VHS cassettes, the first sale doctrine both fostered a new business model and brought an end to the tug of war between studios and rental shops.

It may often seem that the entertainment industries suffer from an inordinate lack of foresight (their resistance to change is certainly a popular subject of lament104), but the movie and music industry are no more shortsighted than anyone else with an established way of doing business. History is replete with stories of technological giants first snubbing, and then

104 See generally JESSICA LITMAN, DIGITAL COPYRIGHT (Prometheus Books, 2006).
being dwarfed by, the upstart mouse.\textsuperscript{105} We all become hostages of our own past success, because we learn the lessons from it \textit{too well}. As a result, “truly innovative start-ups are often the result of creative experimentation with new ideas by outsiders to an industry. … [I]ndifference to industry routines and norms gives an outsider the freedom to break free of the cognitive constraints on incumbents.”\textsuperscript{106} Incumbents simply fail to identify many of the opportunities that are obvious to entrepreneurs, and they fail to recognize the value of those opportunities even after they are made aware of them. And even the innovators can be shortsighted. Bill Gate’s oft-quoted statement: “640K ought to be enough for anybody,”\textsuperscript{107} seems almost apocryphal with the benefit of hindsight, but there are many such examples of innovation flummoxing even the most farsighted of innovators.\textsuperscript{108} Furthermore, there is ample evidence that, once they gain a certain amount of commercial success, innovators themselves tend to stop innovating: “The transition from entrepreneur to established, cautious firm can be breathtakingly fast.”\textsuperscript{109} Of course, the \textit{numerus clausus} provides only a partial remedy for this natural tendency towards innovative inertia, but by limiting the need for users of assets (both tangible and intangible) to get permission to experiment, modify and innovate, it can nonetheless be a powerful antidote to the learned myopia of incumbents.

**The Impact of Market and Social Norms on the Willingness to Freely Give Permission**

Before we leave the issue of the supply side of permissions, we need to say something about the discontinuous effect that price has on the willingness to deal. It might be expected that an offer of a higher price for an asset is more likely to result in a successful transaction than an offer of a lower price. And certainly, this is most often true. It would also seem to follow from this general truism that an offer of a small, albeit positive, price is more likely to

\textsuperscript{105} See examples in J\textsc{ohnathan} Z\textsc{ittrain}, The \textsc{f}uture \textsc{of} The \textsc{i}nternet - And \textsc{h}ow \textsc{t}o \textsc{s}top \textsc{i}t (Yale University Press, 2008).

\textsuperscript{106} Howard Aldrich & Amy Kenworthy, \textit{The Accidental Entrepreneur: Campbellian Antimonies and Organizational Foundings}, in \textit{V\textsc{i}ariations in \textsc{o}rganization \textsc{s}cience: \textsc{i}n \textsc{h}onor of \textsc{d}onald t. \textsc{c}ampbell 20} (Joel A.C. Baum & Bill McKelvey, eds., 1999). For more about the tendency to focus on past experience, \textsuperscript{see R\textsc{ichard} R. N\textsc{elson} & S\textsc{idney} G. W\textsc{inter}, An \textsc{e}volutionary \textsc{t}heory \textsc{o}f \textsc{e}conomic \textsc{c}hange (Belknap Press 1982).

\textsuperscript{107} Bill Gates, 1981, as quoted in G\textsc{raham} T\textsc{anaka}, Digital Deflation 40 (2004).

\textsuperscript{108} “There is no reason anyone would want a computer in their home.” Ken Olson, President, Digital Equipment Corporation, 1977. \textit{Id.} “I think there is a world market for maybe five computers.” Thomas Watson, Chairman of IBM, 1943. \textit{Id.}

result in a successful deal than an offer of a price of zero. However, this is manifestly not true. In fact, it has been shown from behavioral experiments that individuals are more willing to perform a task for free than for a below market rate price, as long as that task is couched in terms of social (rather than market) norms. So, for example, individuals who are willing to work for a charity for free are often not willing to perform the same tasks for a small stipend. In fact, they may even take offense when such amounts of money are given to them. As a result of this dynamic, the effort that individuals will put into an activity, relative to the effort they will put in when they receive no money at all, often decreases sharply. In other words, people are often willing to work for free, but when money is introduced into an activity, individuals expect to be paid at rates that reflect market norms. This documented interaction between market and social norms has been best described by Gneezy and Rustichini: “pay enough, or don’t pay at all.”

Although this behavior may at first seem surprising, it actually has a relatively simple explanation: money moves transactions out of the social sphere and into the market sphere (or from the realm of intrinsic motivation to the realm of extrinsic motivation). The best way to illustrate this is with an example. Suppose you have been having a stressful time at work; knowing this, your wife spends all day in the kitchen making your favorite meal. You are so overwhelmed with gratitude, and filled with delight at your good fortune in choosing such a mate, that you want to thank her in the best way you know how. Now think how offended she would be if you stood up from the table, pulled out your wallet, and offered to pay her for such a wonderful meal. Money moves us out of the social and into the market, and we must choose between the two: they cannot coexist.

But this is not just a laboratory phenomenon, nor is it one that plays out only in the obviously social context of events like a home-cooked dinner. Consider what happened to one consumer, Laura Flores, when she requested permission to resell her copy of a software program, Livingsoft’s ‘Dress Shop 5 Pro,’ after she discovered that there was a restrictive clause in the software’s license agreement. In an email that Livingsoft sent to Laura Flores,

---

112 This example is based on the entertaining examples in Dan Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions* (HarperCollins 2008).
113 Or, even worse, for what would have otherwise come after the meal.
the company stated that it required customers to request permission to resell; but it also required more than this. Livingsoft stated that they gave permissions only when the user had "a financial or physical necessity to give up use of Dress Shop [and when they requested permission] courteously, prior to attempting to sell their software, acknowledging that they are requesting a favor rather than demanding a right."114 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."115

In the end, Livingsoft refused to give Laura permission to resell the software, even though she no longer wanted to use it, had not made copies of it, and had paid over 400 U.S. dollars for it. It should be noted that Livingsoft had absolutely no interest in charging money in exchange for the right to resell: they gave other consumers such permission for free. Instead, Livingsoft seemed to be demanding some sort of noneconomic payment from consumers, a demand which arose because the company felt that its license terms gave it the right to prevent resale. Other consumers had asked ‘courteously’ to resell the product on eBay, but Laura Flores had first attempted to resell without asking for permission, and had later complained about the product online. The other consumers, by acknowledging the rights of the company and framing their requests as ‘favors’, may have tapped into the realm of the social, rather than the market. Because Livingsoft had no interest in charging for permissions, it might have been particularly receptive to this dynamic.

A similar impulse may have been at work in the case of the Wright brothers, or at least in the case of Wilbur Wright. The Wright brothers were notoriously protective of their patent rights, and disputes about their wing-warping patent occupied the courts for years, slowed down the pace of innovation in the field of aviation, and ultimately necessitated the intervention of a government-created patent pool.116 But if we look closer at the historical record, something odd emerges In 1908, in a letter to his brother Orville, Wilbur stated: “If the subscription [to a $100,000 subscription fund set up by the Aero Club of America, whereby its members would help buy the Wright patent] were really spontaneous and of

115 Id.
116 Merges & Nelson, supra note 109 at 890-891.
sufficient size to justify it, I would be willing to make our American patents free." 117 (Emphasis in original.) Despite his notoriously hard-nosed behavior at the bargaining table, Wilbur does seem to be expressing a desire for the sincere and affectionate recognition of his peers, even more than for monetary gain. In effect, the desire for peer recognition is a social norm, which would have moved the licensing of the wing-warping patent from the market sphere into the social sphere.

Although the interaction between market and social norms has not previously been analyzed in terms of an endowment effect, the phenomena do seem to be connected. The moral claim of ownership by definition arises by virtue of property ownership, and it certainly does increase the price demanded for an asset, albeit in terms of a non-monetary payment (which, in some circumstances, may be even less forthcoming from the buyer than more money). In other words, property ownership might cause what might be termed a ‘noneconomic endowment effect’, in that the holding of a right can give rise, not only to barriers to market transactions, but to social transactions as well. If a social acknowledgment of ownership is forthcoming, then property owners sometimes choose to move from the market economy into the gift economy, from the market sphere into the social sphere. But if such acknowledgement is not forthcoming, what would have been given for ‘free’ is then withheld entirely. 118 Indeed, it may be that in some cases, the presence of this effect actually exacerbates the economic endowment effect. That is to say, if a property owner is thwarted in his attempt to frame a transaction as ‘social’ (i.e., as a favor or a gift), he might demand an even higher price in the market sphere (i.e., than he would have demanded if he had simply entered the market from the outset). Something like this certainly seems have been at work in the case of Wilbur Wright, for example. 119

Part 3

DIVIDED RIGHTS VERSUS FRAGMENTED RIGHTS

Most of the changes in behavior that we have examined in the previous two sections are implicated whenever property rights of any type are involved, including those that unify

118 This is similar to what happens when economic incentives are given for charitable goods - donations actually decrease. See e.g. Carl Mellstrom & Magnus Johannesson, Crowding Out in Blood Donation: Was Titmuss Right?, 6 J. EUR. ECON. ASS'N 845 (2008); and Edward L. Deci et al., A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivations, 125 PSYCHOL. BULLETIN 627 (1999).
119 Howard, supra note 117.
rights of use and possession. So, for example, a full owner and possessor of a widget may refuse to sell, to temporarily transfer, or to grant permission for the use of that widget, even if such transactions would make everyone better off. However, the ability of a property owner to ‘hold out’ in the face of a more productive use by another person is the basic cost of having any property rights system at all. For this reason, the law is highly tolerant of fragmented rights – i.e., of fragmentation of rights along the tangible dimension (you can cut a widget in half and sell half a widget). On the other hand, the law is highly intolerant of divided rights – i.e., of divisions or separations of use and possession (the law would likely interfere if you sold a widget subject to a use restriction).

The second reason why it makes sense for the law to treat divided rights and fragmented rights differently is that there is a productive advantage to unifying rights of use and possession. We know that users have advantages in seeking out innovations that can be revealed only through asset use, and that the innovations that arise out of asset use are economically valuable. Therefore, resource values will be higher when rights of use and possession in assets are unified. The positive productive impact of unifying rights of use and possession can be distinguished from the more general problem of matching property rights inputs and human capital inputs. The matching problem will exist under any property rights allocation: we can never be completely sure that the ‘right’ stock of human capital is matched with the ‘right’ asset. However, we can say that once physical possession of an asset has already been transferred, it is beneficial if use rights in that asset transfer along with it.

The third reason why it makes sense for the law to treat divided rights and fragmented rights differently is that feelings of psychological ownership have particularly pernicious effects when rights of use and possession are divided - effects that are not present when use and possession are unified. We would expect that use and possession of half a widget would give rise to feelings of psychological ownership of just that – half a widget. Thus, psychological ownership would not itself interfere with a transaction requiring the combination of two halves of a widget. Furthermore, two owners of half a widget each may

---

120 It has been noted in the expropriation context, for example, that unrestrained application of the principle of moving a resource to its higher value use deprives property rights of all force and meaning. See e.g. Walter Block, Coase and Kelo: Ominous Parallels and Reply to Lott on Rothbard on Coase, 27 WHITTIER L. REV. 997 (2006) [Block].
121 This was discussed in detail in Chapter 2.
122 ‘Right’ is defined in terms of maximizing innovation and value.
be forced to deal with each other simply because of the physical nature of the assets: if only a full widget is functional, then physical necessity provides a powerful incentive to reach a successful deal.

The situation is quite different when it involves a possessor of a full widget and an owner of use rights in that widget. In this scenario, both possession and legal ownership give rise to feelings of psychological ownership. The best way for the owner of the use right to express those feelings may be via the denial of permission. The best way for the possessor of the asset to express those feelings is to use the asset without asking for permission at all. And since the possessor already has the full widget, in most cases he will simply be able to unilaterally violate the use right owner’s legal rights (this violation, in turn, creates a powerful incentive to hide the use). Together, all of these factors mean that barriers to recontracting are especially high when rights of use are divided from the right of possession. Unifying rights of use and possession completely eliminates some of these barriers to recontracting, and greatly reduces others. Thus, we can see that the user innovation theory of the *numerus clausus* fully accounts for the law’s differential treatment of fragmented and divided rights: it makes sense for the law to treat divided rights and fragmented rights differently because feelings of psychological ownership operate quite differently in these two cases.

**Part 4  PSYCHOLOGICAL OWNERSHIP VERSUS ANTICOMMONS THEORY**

Before we conclude our discussion of how psychological ownership interferes with *ex post* recontracting, we need to examine how its effects differ from a body of economic theory which concerns itself with fragmented (rather than divided) rights. In Chapter 1, we showed that the economic theories of the *numerus clausus* do not adequately account for the law’s preference for exclusion. But there is also another body of research which has a related set of concerns: anticommons theory. Anticommons theory, developed most extensively by the work of Michael Heller, is built around the notion that multiple rights of exclusion (i.e., an anticommons) can result in underuse of a resource, just as multiple rights of use (i.e., a
commons) can result in overuse and resource depletion.\textsuperscript{123} Thus, just as we have the well-known ‘tragedy of the commons,’ we have the lesser-known ‘tragedy of the anticommons.’\textsuperscript{124} An anticommons can be created in a variety of ways: when multiple individuals have full ownership rights in a single asset (A and B are joint owners of a widget); when single individuals each have full ownership rights in separate assets which must be combined (A owns half a widget and B owns the other half); and when rights of use and possession in a single asset are divided across different individuals (A owns the right to use the widget and B has the right of possession but not use).

Although anticommons theory has proven to be an important and fruitful field of research, it does not adequately explain the law’s preference for exclusion. As Merrill and Smith note, the law does not police the number of owners of an asset: “The primary candidate for an economic explanation has been the suggestion that the \textit{numerus clausus} is a device for minimizing the effects of durable property interests on those dealing with assets in the future, and in particular the effects of excessive fragmentation of interests, or an "anticommons." On this view, the \textit{numerus clausus} serves to prevent situations in which too many individuals have a veto right over the use or disposition of a resource. But whatever the merits of this anti-fragmentarian view for other property doctrines, it does not fully explain the \textit{numerus clausus}, which is aimed at limiting types of rights, not the number of rightholders.”\textsuperscript{125}

More recently, a group of scholars, spearheaded by Francesco Parisi, have combined anticommons theory with price theory to explain what they see as the law’s aversion to ‘fragmented property.’\textsuperscript{126} These theorists suggest that even when the costs of recombining fragmented property is zero, individuals will not recombine rights because each individual

\begin{footnotesize}
\textsuperscript{126} See generally, Norbert Shulz, Francesco Parisi and Ben Depoorter, \textit{Fragmentation in Property: Towards a General Model}, 158 J. INSTITUTIONAL & THEORETICAL ECON. 594 (2002) [Fragmentation]. Fragmented property is defined to include all three of our widget scenarios outlined above, as long as there is no rule on sharing the proceeds of a sale in any of the scenarios. (This is because the rule for sharing the proceeds gets rid of the dispute about how to divide the value of the property.)
\end{footnotesize}
will seek to appropriate the full value of the recombined asset.127 This theory predicts that the
tendency towards holdouts will increase with the degree of fragmentation, the
complementarity of assets, and the independence of the pricing decisions.128

However, the combination of price theory and anticommons theory does not prescribe
legal outcomes that coincide with the law’s approach to fragmented property. The price
theory/anticommons rationale is strongest in the case of land (where inputs tend to be
complements, rather than substitutes), and weakest in the area of chattels (where there are
usually many substitutes available).129 In contrast, the law’s concern about the separation of
use and possession is weakest for land, and strongest for chattels. In addition, the law is
highly averse to even simple, bilateral separations of use and possession, which should
produce no more troublesome hold out behavior than mere property rights ownership itself.

In addition to these theoretical and predictive mismatches with the law, neither
anticommons nor price/anticommons theory explains why the law is highly tolerant of shared
ownership (multiple owners of a single widget), and fragmented rights (you can cut a widget
in half and sell half a widget), but is highly intolerant of divided rights - divisions or
separations of use and possession (the law would likely interfere if you sold a widget subject
to a use restriction). In fact, both versions of anticommons theory lump these three scenarios
together and treat them alike.130

Anticommons theory is not alone in equating fragmented rights (or fragmented
assets), and divisions of use and possession (divided rights). Hansmann and Kraakman, for
example, do the same in their paper on the numerus clausus: “To say that two persons hold
divided rights in the same asset is little different from saying that the two persons own
undivided rights in distinct assets.”131 They assert that it is possible to conceptualize
fragmented rights and divided rights as equivalent, simply by redefining the ‘asset’: “Suppose, for example, that A, who owns an acre of land in fee simple, divides the land into

127 Id. at 597.
129 Parisi et al assert that neither strict complementarity, nor a high level of fragmentation is required by their theory. See Fragmentation, supra note 126 at 597, 605. However, if we interpret even bilateral divisions of rights as ‘anticommons,’ then all private property rights become problematic. Indeed, in such a framework, private property rights solve no allocational problems at all.
130 See generally, Fragmentation, supra note 126.
two adjoining half-acre parcels and sells one of the parcels in fee simple to B. Do A and B now each have undivided ownership of distinct assets? Or do they have divided ownership rights in a single asset, the given acre of land? The law and ordinary language both generally conceptualize the situation in the former fashion. But… it could equally well be viewed the other way.”132 They also assert that the same is true if we divide the parcel of land along its temporal or its use dimension, rather than its physical dimension. Thus, if we convey the use of the entire parcel to A on Mondays to Wednesdays, and to B on every other day, we can either say that A and B share ownership of the entire parcel, or that A and B own different assets (where the assets are now defined as ‘the land on Mondays to Wednesdays’ and ‘the land on Thursdays to Sundays’).133

Although legal and economic scholars conceptualize fragmented rights and divided rights as equivalent, they are not equivalent from the point of view of psychological ownership. Consider the case of a physically divided widget. We would expect that use and possession of half a widget would give rise to feelings of psychological ownership of just that – half a widget. Thus, psychological ownership would not interfere with a transaction requiring the combination of two halves of a widget. In a shared or multiple ownership scenario, where more than one individual shares ownership of a single asset, feelings of psychological ownership once again coincide with legal ownership. It is only in the case of a separation of use and possession that there is a mismatch between psychological ownership and legal ownership, because both possession and legal ownership give rise to feelings of psychological ownership. Of course, both the shared ownership and fragmented rights scenarios may give rise to disputes over value, but the common law of property (at least in regards to the numerus clausus134) does not concern itself with disputes over value. Instead, it only concerns itself with disputes over ownership. We will discuss this proposition in the next section.

132 Id. Hansmann and Kraakman do not seem to consider whether these two scenarios are different from a third scenario – the true multiple ownership scenario, where A and B are joint owners of the entire parcel.
133 This example is adapted from Hansmann & Kraakman. Id.
134 The rules of joint ownership may, for example, be animated by a concern for disputes over value. See Ariel Katz, Copyright Collectives: Good Solution But for Which Problem? in WORKING WITHIN THE BOUNDARIES OF INTELLECTUAL PROPERTY LAW (Harry First, Rochelle Dreyfuss, Diane Zimmerman, eds., 2010). Available at SSRN: http://ssrn.com/abstract=1416798.
Value Versus Ownership

Anticommons theory posits that when complementary assets have to be combined in order to realize a productive use, holdouts will occur because each individual owner will seek to appropriate the full value of the combined assets. This may no doubt be true, and it may provide justification for legal intervention in exceptional cases of holdout behavior - when we think that expropriation is warranted because of the public value of the resulting use. However, anticommons theory does not explain the *numerus clausus*, because the *numerus clausus* is about disputes over ownership, not disputes over value. This is not to say that anticommons theory does not explain disputes over value – in fact, it does this quite well. However, the *numerus clausus* seeks to prevent mismatches between legal and psychological ownership - by interfering with the legal enforcement of use restrictions on assets. In contrast, the *numerus clausus* does nothing to prevent or mitigate disputes over value, because it does not interfere when tangible assets are divided into many fragments and each fragment is conveyed to a separate owner.

The observation that the *numerus clausus* (and the common law of property more generally) concerns itself with disputes over ownership (but not disputes over value) marks an important point of departure between psychological ownership and the concerns that animate anticommons theory, and lends descriptive as well as normative force to the user innovation theory of the *numerus clausus*. However, the relevance of the value/ownership distinction goes beyond the user innovation theory of the *numerus clausus*, and has implications for the coherence of private property more generally. By structuring itself to reduce disputes over ownership rather than disputes over value, the common law prevents the concept of private property from becoming entirely meaningless. In other words, if disputes about value invited legal redress every time a property owner decided to hold out for a bigger share of the pie (or because of an absolute mismatch between social values and private values), then there is no such thing as private property. That is, if we legally force a reallocation of property every time a higher value user comes along, then there is simply no

---

135 The fact that anticommons theory does not explain the *numerus clausus* has also been noted by Merrill & Smith, *supra* note 125.
136 A more detailed discussion of how psychological ownership explains the *numerus clausus’* treatment of tangible property can be found in Chapter 4, Part 2, *Validating Psychological Ownership*.
private property at all, because the basic right of private property is the ability to hold out in
the face of a more productive use by a non-owner.\footnote{137}

It is true that the most rhetorically forceful version of anticommons theory is
expressed as being about ‘too much property,’ because it sounds laudable to ensure that
property is not subject to too many greedy (or misguided) hands. However, in the case of
tangible assets, a highly fragmented anticommons scenario is little different from a simple
misallocation of rights. That is, if a developer needs a particular parcel of land in order to
realize a higher value use, and the owner(s) refuse to sell, it does not matter if that parcel is
fragmented into 10 pieces, 2 pieces, or whether the parcel is simply owned by a single
individual who refuses to sell. And indeed, Parisi et al admit that anticommons theory is fully
applicable to the simple case of resale of a single, unfragmented asset.\footnote{138} In other words,
when applied to tangible assets, anticommons theory really boils down to a judgment that
‘the wrong person owns the asset.’ Psychological ownership rescues the law from this
dilemma, because it only applies when use rights are separated from permanent possession
(that is, when possession has already been permanently and voluntarily transferred).\footnote{139} Thus,
psychological ownership provides the law with an internal limit which preserves the concept
of private property.

This is not to say, however, that anticommons theory has nothing to important to tell
us, because there is another area of property law which cares very much about the
apportionment of value – intellectual property law. Even though the U.S. Constitution
grounds intellectual property protection squarely within an incentive-based, public purpose
framework,\footnote{140} the role of natural law and Lockean impulses in intellectual property law
cannot be denied.\footnote{141} And whether or not we formally conceive of intellectual property law as
being restitutionary in nature, it is nonetheless true that the animating force in intellectual
property law is the tension between rewarding the contributions of sequential innovators,

\footnotesize{\footnote{137} When we expropriate property, the ‘public use’ criterion obscures this fact, but the trend towards considering
increased tax revenue as a public use makes it more obvious. This is one reason why Kelo v. City of New
London, 545 U.S. 469 (2005) was so controversial.}

\footnotesize{\footnote{138} Although they assert that it would only apply to the resale of a single undivided asset, rather than to the
original sale of that asset. See Francesco Parisi, \textit{Entropy in Property}, 50 Am. J. Comp. L. 595, n. 4 (2002):“even
reversing a simple property transaction can result in monopoly pricing by the buyer-turned-seller.”}

\footnotesize{\footnote{139} See Chapter 4, Part 2, 2) \textit{Possession and the Duty to Return.}}

\footnotesize{\footnote{140} U.S. Const. Art. 1, § 8, cl. 8.}

\footnotesize{\footnote{141} See e.g., Wendy Gordon, \textit{On Owning Information: Intellectual Property and The Restitutionary Impulse}, 78
Va. L. Rev. 149 (1992).}
While still allowing innovation to occur. Thus, intellectual property law can be understood as a continuing compromise between the restitutioary claims of successive generations of creators.

In addition to this underlying concern with apportioning value to creators, there is an even more important reason why anticommons theory has more force in the intellectual property arena. In the case of nonrival intangible assets, preventing anticommons problems involves the denial of intellectual property protection altogether. When we deny intellectual property protection to an asset, we do not transfer the asset from one private party to another. Instead, the asset is free for all to use. For this reason, anticommons theory retains its persuasiveness when applied to intellectual property. However, we should be cautious about taking the anticommons view seriously in the case of tangibles, because the application of the anticommons perspective to tangible assets simply transfers wealth from one private party to another.142

CONCLUSION

In this Chapter, we saw that even when the traditional transaction costs of measurement and enforcement are low, intransigent behavioral barriers can prevent individuals from contracting around governance rules. These behavioral barriers are particularly robust when rights of use in an asset are separated from the right to its possession.

Possession of an asset, and legal ownership of use rights, both give rise to feelings of psychological ownership. Therefore, when rights of use are divided from the right of possession, both the possessor of the asset and the owner of the use rights feel that they own the entire asset. Individuals will not contract over rights that they feel they already ‘own’, and so the fluid process of recontracting, to which clear notice of use restrictions should give rise, does not occur. Thus, recontacting is stymied, not just by the obvious costs of contacting,143 but even more so by an unwillingness to recontact that arises from the

---

142 This observation has been made by other scholars with respect to transaction cost economics more generally (of which anticommons theory is a subset). See generally, Block, supra note 120.
143 Contacting the rights-holder, negotiating the price and scope of permissions, etc.
divisions of rights themselves. This unwillingness to recontract around divided rights means that governance rules will not produce the same innovative outcomes as rules of exclusion. Because rules of exclusion give users full use rights in the assets that they possess, exclusion gives users the freedom to innovate without having to get permission from the owners of use rights. In this way, exclusion entirely circumvents the recontracting problem.

Now that we have examined the impact that psychological ownership has on contracting behavior, it is appropriate to examine whether psychological ownership (and user innovation) can account for the detailed contours of the *numerus clausus*.
Chapter 4: The User Innovation Theory and Tangible Property

“Every assignee of a lease has notice of the lessor's covenants; consequently no covenant, however absurd soever, 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.”¹

INTRODUCTION

In Chapter 1, we saw that the numerus clausus is the law’s way of implementing its preference for exclusion (which also entails limiting the proliferation of governance rules). This legal preference provided us with a puzzle in need of theoretical justification. In Chapter 2, we began our search for that justification by examining the critical role that exclusion plays in fostering user innovation. Thus, we saw that the law should promote user innovation because it directly increases resource values. In Chapter 3, we examined the question of why governance rules will not produce innovative outcomes equivalent to those produced under rules of exclusion. There, we saw that the recontracting process will not result in the rearrangement of rights that is necessary to foster user innovation because feelings of psychological ownership interfere with the recontracting process. Thus, we saw that the law should validate psychological ownership by bundling rights of use with the right of possession.

Together, Chapters 2 and 3 have thus provided us with a general economic rationale for the law’s preference for exclusion. Now, it is appropriate to examine whether the dual guideposts identified in Chapters 2 and 3 - user innovation and psychological ownership - can account in a more fine-grained way for the specific legal doctrines via which the law implements its preference for exclusion vis-à-vis tangible assets. For example, can the law’s resistance to the creation of novel servitudes be explained in terms of the promotion of user innovation, or the validation of psychological ownership? Thus, this Chapter answers the question of whether (and how) the specific legal doctrines that implement the numerus clausus promote user innovation or validate psychological ownership.

This Chapter shows how the user innovation theory of the numerus clausus accounts for several features of the law of property which have long puzzled judges and scholars. In Part 1, we examine how user innovation explains the following features of the numerus clausus: 1) the legal solicitude for the right to full use (which is expressed via the legal restrictions on the creation of novel servitudes); and 2) the differential legal treatment of servitudes on land and servitudes on chattels. In Part 2, we see how psychological ownership explains the following features of the numerus clausus: 1) the law’s tolerance for fragmented

---

2 We will deal with intangible assets in Chapter 5.
rights, but intolerance for divided rights; 2) the law’s refusal to recognize permanent leases; and 3) the fact that notice alone does not determine the boundaries of the *numerus clausus*.

**Part 1  PROMOTING USER INNOVATION**

To see how specific features of the *numerus clausus* promote user innovation, let us recall our three assertions about user innovation:³

1) Users of assets will have an advantage in *problem finding*⁴ because:
   i) asset use will tend to reveal problems; and  
   ii) differential preferences will create different patterns of asset use and so reveal different sets of problems.⁵

2) Users of assets will have an advantage in *problem solving* because:
   i) asset use provides a low-cost, repeated, trial and error process;  
   ii) differential cognition across users will lead users to perceive different sets of possible solutions to problems; and  
   iii) 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.⁶

3) 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.⁷

The first two assertions establish two important facts about user innovation: 1) that users will have advantages in seeking out new uses of assets, advantages that cannot be duplicated by non-users; and 2) that each user is likely to innovate in a unique manner, and so one user is not a substitute for another user. The third assertion points to the role that full

---

³ These assertions were discussed and developed in Chapters 2 and 3.  
⁴ The difference between problem finding and problem solving was discussed in Chapter 2, *Problem Finding and Problem Solving Contrasted*.  
⁵ See Chapter 2, *Problem Finding*.  
⁶ See Chapter 2, *Problem Solving*.  
⁷ See Chapter 3, *The Motivation to Disclose*.  

135
use of an asset plays in the disclosure of user innovation. Together, these three assertions indicate that full use rights should accompany the permanent possession of assets whenever it is desirable to have a robust search for new uses of assets.\(^8\)

The law expresses its solicitude for unifying the full use of an asset with the possession of that asset (but only when it is desirable to have a robust search for new uses of assets) in two main ways. First, the law generally refuses to enforce governance rules that encroach upon the right to reuse\(^9\) an asset - these are the restrictions on the creation of novel servitudes. Second, the law has a greater tolerance for servitudes on land than for servitudes on chattels. We will discuss each of these in turn.

1) The Legal Policy Against the Creation of Novel Servitudes

When a seller wishes to restrict the downstream uses of an asset, contract provides an obvious way to accomplish this vis-à-vis the original seller and the initial buyer. However, downstream transferees lack contractual privity with the original seller, and the cumbersome device of assigning contractual obligations is risky.\(^10\) Thus, a seller who wishes to restrict the uses of an asset no matter how many times that asset changes hands will want to create a right that is enforceable against the world – i.e., a property right. Of course, a seller who wishes to create such a right is faced with the pronouncement in *Keppell v. Bailey:* “‘[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 any owner.’”\(^11\) The common law restricts the creation of use restrictions on both land and chattels. However, because the rules that apply to each of these categories differ markedly, we will discuss each type of asset separately.

---

\(^8\) We will discuss when it is desirable to have a robust search for new uses of assets *infra,* The Law has a Greater Tolerance for Servitudes on Land than for Servitudes on Chattels.

\(^9\) The use of the term ‘right’ here is not correct in the strict Hohfeldian sense, which would be: the power to resell, and the privilege to full use.


\(^11\) *Keppell v. Bailey,* *supra* note 1 at 1049.
i) Land

Even though land use restrictions represent a marked departure from the policy against restraints on alienation - which strongly resists the ability of former land owners to enforce use restrictions on land parcels that have been transferred to a new owner - land use restrictions have long existed in the common law. There are four general types of land use restrictions: easements, profits, real covenants and equitable servitudes. Together, these are generally referred to as servitudes. An easement is a nonpossessory right to enter upon or use the land of another. Profits, the right to extract natural resources from another’s land, are closely related to easements; indeed, an easement is often required for the right of profit to be exercised. There are differences between the two rights, however. For example, at English law, profits can be held in gross, while easements cannot.

Real covenants are promises which relate to the use of land. Unlike easements, which give an individual a right of enjoyment in another’s land, covenants regulate what an owner can do on his own land. While these promises can always be enforced between the original promisee and promisor as a matter of contract law, the question of when these promises can be enforced vis-à-vis subsequent transferees of the land has generated a complex set of rules in both law and equity. At law, the benefit of a real covenant only ‘runs with the land’ if certain conditions are met, while the burden does not run at all. Real

---

13 Because of the similarity between profits and easements, they are sometimes considered together. So, most American scholars divide them into three categories: easements, real covenants and equitable servitudes. See e.g. Id. at 1261.
14 Id.
17 Id. at 113 and 262. Easements, in other words, require that there be both a dominant and servient tenement. In addition to the requirement of a dominant and servient tenement, other requirements for easements are: the easement must accommodate the dominant tenement, the dominant and servient tenement must have different owners, and the easements ‘must be capable of forming the subject matter of a grant’. See In Re Ellenborough Park [1956] Ch. 131 at 163. An easement is capable of forming the subject matter of a grant if: 1) the rights are not too wide and too vague in character; 2) the rights are not rights of mere recreation; and 3) the rights do not deprive the owner of the servient tenement of possession. In Re Ellenborough Park [1956] Ch. 131 at 175-6.
18 Rose, Comments, supra note 15 at 1405.
19 See e.g. French, Reweaving, supra note 12 at 1269-1276 (American law); and Law Commission, supra note 16 at 126-131 (English law).
20 That the burden did not run at law is well established: Keppell v. Bailey 39 Eng. Rep. 1042 (Ch. 1834); Austerberry v. Oldham (1885) 29 Ch. D 750; Rhone v. Stephens [1994] 2 AC 310.
covenants are only transformed into property interests that ‘run with the land’ (i.e., which are enforceable against subsequent transferees) if they ‘touch and concern’ the land to be benefited, and if there is ‘vertical privity.’

The restrictions on the running of the benefit, along with the fact that the burden did not run at law, meant that the land use plans of private owners were frequently thwarted by the courts. Beginning in *Tulk v. Moxhay*, the English courts responded to the private demand for more flexibility in the creation of land servitudes by resorting to equity to loosen the strictures of the *numerus clausus*. In that case, the court decided that the burden of a covenant would run if the subsequent transferee of the land had notice of it. However, even at its most activist, the judiciary was not without serious reservations about the ability of servitudes to encroach upon the rights of property owners. As a result of this discomfort, *Tulk v. Moxhay* was narrowed over time by the introduction of additional limitations for the burden to run in equity: 1) the covenant must be restrictive, rather than affirmative; 2) the covenant must ‘touch and concern’ the benefited land; and 3) the original covenanting parties must have intended that the burden run with the land.

American courts have been even more adventurous than the English courts, expanding the availability of land servitudes in several ways. Easements in gross, for example, continue to run afoul of the law in England, but they are now permissible in most

---

21 Vertical privity refers to the requirement that both the original covenantee and his successor in title had a legal estate in the benefited land. See e.g. Susan French, *Highlights of the New Restatement (Third) of Property: Servitudes*, 35 REAL PROP. PROB. & TR. J. 225, 233-234 (2000) (vertical privity) [French, Highlights]. There is another class of real covenants, those between landlord and tenant, for which horizontal privity is required; however, this class of covenants will not be discussed here. Id. at 229-230 (horizontal privity).

22 41 Eng. Rep. 1143 (Ch. 1848).

23 “It is said that, the covenant being one which does not run with the land, this court cannot enforce it, but the question is not whether the covenant runs with the land, but whether a party shall be permitted to use the land in a manner inconsistent with the contract entered into by his vendor, with notice of which he purchased.” *Id.*

24 There were also rules for when the benefit will run at equity in English law, which have now been largely replaced by statutory annexation. See *Federated Home v. Hill Lodge Properties* [1980] 1 WLR 594.

25 *Austerberry v. Oldham* (1885) 29 Ch. D 750.


27 See e.g. Thomas W. Merrill & Henry E. Smith, *Optimal Standardization in the Law of Property: The Numerus Clausus Principle*, 110 YALE L.J. 1, 17 (2000) (citations removed) [Merrill & Smith]: “The innovation wrought by the Court of Chancery was quickly limited, as subsequent decisions held that equity would enforce promises as property only if the promise benefits an appurtenant interest in land… and only if the promise "touches and concerns" the land.”

Despite the more permissive attitude of the American courts, many of the common law restrictions imported from the English courts continue to exist in some form. The requirements of touch and concern, privity of estate, dominant and servient tenements, among others, all attest to a continuing judicial sense that servitudes cannot be allowed to proliferate unchecked by the interfering hand of the law.

ii) Chattels

The rules concerning the creation of servitudes on chattels are much simpler than those that govern land servitudes. Historically, in both law and equity, there was an absolute bar on the creation of servitudes on chattels: “Conditions of this kind do not run with goods, and cannot be imposed upon them.” But although the policy against giving legal countenance to servitudes on chattels retains much of its historic vigor, American courts loosened its strictures in a variety of cases. These can be divided into three rough categories: 1) the business transaction cases; 2) the coupon and ticket cases; and 3) the product safety and business reputation cases (many of which also involve trademark issues). The business transaction cases involve either the sale of a business which includes servitude-burdened chattels, or the sale of a major moveable business asset, such as a ship or boat.

---

30 For a similar argument, see Law Commission, supra note 16 at 12: “Cases subsequent to Tulk v. Moxhay have reflected this tension between the contractual nature of covenants and their proprietary effect; they affirm the proprietary effect but subject to a number of complex limitations the total effect of which is difficult to justify. It is arguable that some of these difficult rules spring form the discomfort of the courts with the apparent contradiction inherent in the concept of covenants that behave like property rights.”
31 Taddy v. Sterious, [1904] 1 Ch. 354 at 358.
32 There is actually a fourth category, the patent field-of-use and the copyright license cases, though these are not considered to be servitude cases at all. This is because the presence of the intellectual property right allows a court to merely determine what restrictions are within the intellectual property grant, and which fall outside of the grant. See e.g. the discussion infra Chapter 5, Part 1, First Sale.
33 The most famous example of the former is the sale of the Joy Bar Luncheonette in Pratte v. Balatsos, 99 NH 430, 113 A2d 492 (1955). The owner of the restaurant, Albert Larochelle, had signed a contract with Oscar Pratte, whose company leased juke boxes to various businesses, splitting the receipts from the coin operated machines. Among other things, the contract between Larochelle and Pratte obligated Larochelle to maintain the juke box in the restaurant for fourteen years and six months, and to bind his heirs, successors and assign assume all obligations of the contract. When Larochelle later sold the luncheonette to Spiros Balatsos, he told Spiros about the juke box contract, but there was no assignment of the contractual obligations. Balatsos then tried to force Pratte to remove the juke box, but the Supreme Court of New Hampshire held that Spiros’ knowledge of the obligations between Larochelle and Pratte created an equitable servitude on the juke box. Thus, he could neither remove the machine, nor unilaterally alter the terms of the sharing agreement.
34 The latter type of business transaction case, which involves the sale of a chattel that constituted the main part of a business enterprise, is exemplified by the marine vessel cases. A typical example of this type of case is Tri-
The business transaction cases impose quite onerous duties on the purchasers of chattels.\textsuperscript{35} However, they at least involve a limited number of assets and an actual course of negotiation between the parties. Thus, their overall impact on user innovation is low.\textsuperscript{36}

Of the various types of chattel servitude cases, the coupon and ticket cases are among the most likely to result in legal enforcement of the servitude. These cases usually involve the enforcement of non-transferability clauses against downstream transferees of items like airline and train tickets, and coupons which are redeemable, or give fare reductions, for such items. Because these items are essentially future service contracts which include a permission to trespass on a business premises, the ongoing burden on the service provider can be said to justify the non-transferability in these cases. And perhaps for this reason, these cases are often resolved in terms of contractual principles, even though courts usually take notice of the servitude problem in their midst:

“Thus, for purposes of the case before us, the question of whether TWA's tariffs may be enforced seems to be a question of whether its frequent flyer coupons embody rights of property or of contract. Unfortunately, this bare dichotomy is not very helpful without more. "Property rights" and "contract rights" do not have independent existence in the world as natural kinds, detached from any consideration of human purposes. Nor are the categories mutually exclusive in common usage: When property changes hands, it quite often does so pursuant to a contract (as in this case); and contract rights, having an economic value (again, as in this case), are often referred to as "property," as they surely are for some purposes. … In any event, although we find that prior examinations of the conflicting concepts of property and contract rights do not provide entirely satisfactory answers to the problem that confronts us, we believe that the weight of authority strongly supports the conclusion that it is the latter and not the former that are at issue here. Although there are no other appellate decisions addressing the character of the rights embodied by frequent flyer coupons, decisions in analogous areas suggest that these innovations in the travel market evidence rights of contract, Continental v. Tropical Marine Enterprises, 265 F2d 619 (5th Cir 1959), where a mortgagee was found to be bound by restrictions on the use of a vessel in specified ports because they "r[a]n with the boat" as an equitable servitude (at 623). Thus, the mortgagee, in foreclosing on the vessel, had to make the subsequent sale of the vessel subject to the port restrictions, even though it had never specifically contracted to do so.

\textsuperscript{35} Even in the case of a seemingly innocuous item like a juke box, we can question the burden of maintaining the use of a chattel when it might interfere with the running of a business. See e.g. Zechariah Chafee, Jr., The Music Goes Round and Round: Equitable Servitudes and Chattels, 69 HARV. L REV. 1250, 1258 (1956).

\textsuperscript{36} See discussion infra Part 2, User Innovation and the Mechanisms of Governance.
and, thus, that the public policy against restraints on the alienation of property is no impediment to the enforcement of TWA's tariffs.”

The third category, the product safety and business reputation cases, also result in a relatively high level of servitude enforcement. As the name suggests, in this type of case, the asserted purpose of the post-sale restriction on the chattel is the protection of consumers from defective or dangerous products, and the prevention of the related loss of business reputation and goodwill. However, some of them are enforced on the basis of an explicit chattel servitude rationale, where the business goodwill is the dominant tenement and the product is the servient tenement: “If equitable servitudes on chattels can be sustained on the theory that the producer has a proprietary interest therein for the benefit of his business, why not the purchaser-retailer of damaged goods who also has his own good will to maintain?”

Hair coloring products in particular seem to generate a disproportionate degree of servitude enforcement. The servitudes in these cases generally involve the enforcement of ‘professional use only’ labels to prevent the resale of lower priced hair color from professionals to home users. Although these resale restrictions essentially protect a scheme of market segmentation, most of the hair color cases involve a mixture of both consumer protection and business goodwill rationales, because professional use only hair color is sold without warning labels or instruction sheets. Although some courts enforce resale restrictions on hair color, others opt for a more targeted approach, and hold that as long as an information sheet is attached to the product, reselling hair color packaged for beauty salons is perfectly acceptable.

---

38 In one of the trading stamp cases, the court was even more clever in its circumvention of the chattel servitude problem. It decided that: “The right to redeem the stamps is a property right transferable by possession while the license to use them for advertising purposes is not transferable without compensation to the person granting that right.” Sperry & Hutchinon Co v. Fenster, 219 F 755, 757 (ED NY 1915).
40 Individual manufacturers usually sell the same chemical product to beauty salons and to consumers, but with different packaging and instructions sheets. Beauty salons are sold lower-priced hair color in bulk, without individual boxes and instruction sheets; while consumers are sold higher-priced hair color in individual boxes, with product warnings and detailed instructions on how to use the product.
41 Nonprofessional users may suffer poor results, hair breakage, or even bodily injury if they use the product improperly. In addition, selling unboxed hair color exposes the product to light, and this can reduce its effectiveness. Both of these can have a negative impact on product reputation, because the consumer is likely to blame the manufacturer, rather than the enterprising reseller.
iii) Scholarly Criticism of the Legal Policy Against Servitudes

Although American courts have opened up the availability of both chattel and land servitudes, there remains a stark mismatch between the selective and relatively cautious approach of the judiciary, and the attitude of those legal scholars who write in support of the greater availability of servitudes.\(^{43}\) Indeed, the latest Restatement\(^{44}\) has advocated eliminating or replacing almost all of the doctrines that currently restrict the availability of servitudes: horizontal privity, limitations on affirmative burdens, limits on benefits in gross, limits on third party beneficiaries, the touch and concern requirement and vertical privity.\(^{45}\)

The ‘pro-servitude’ scholarship puts forward five principal arguments in favor of abandoning the legal restrictions on the creation of novel servitudes.\(^{46}\) The first argument is that the rules restricting the availability of servitudes, both in law and in equity, are both complex and outdated: “The law of easements, real covenants and equitable servitudes is the most complex and archaic body of American property law remaining in the twentieth century.”\(^{47}\) A second and related argument is that servitude law is unfair because this complex set of rules can be circumvented via clever legal gymnastics: “actors who are sophisticated or well-advised can almost always manipulate the menu of options so as to realize their objectives. In this sense, the *numerus clausus* discriminates in favor of those who are well-endowed with legal resources and against those who are poorly endowed.”\(^{48}\)

The third argument against servitude restrictions is that freedom of contract demands that we allow private parties to divide their property rights as they wish: “with notice secured by recordation, freedom of contract should control.”\(^{49}\) According to this view, property rights

---


\(^{44}\) Restatement (Third) of Property: Servitudes (2000).

\(^{45}\) For an explanation of these proposed changes, see generally French, Highlights,* supra* note 21.

\(^{46}\) There is also plenty of scholarship on the other side of the debate, of course. See *e.g.* Reichman, Unified,* supra* note 28 (land servitude requirements protect individual liberty and prevent inefficiency); Stewart E. Sterk, *Freedom From Freedom of Contract: The Enduring Value of Servitude Restrictions*, 70 IOWA L. REV. 615 (1985) (land servitude requirements needed because of third party externalities and limited foresight); and James Winokur, *The Mixed Blessings Of Promissory Servitudes: Toward Optimizing Economic Utility, Individual Liberty, And Personal Identity*, 1989 WIS. L. REV. 1 (land servitudes can conflict with utility, individual liberty and personal identity).

\(^{47}\) French, Reweaving, *supra* note 12 at 1261.

\(^{48}\) Merrill & Smith, *supra* note 27 at 7.

\(^{49}\) Epstein, Notice, *supra* note 43 at 1358.
regimes - whether it be the *numerus clausus* or intellectual property laws – are nothing more than default rules which provide a starting point around which contracting can occur: “In transactional relationships, property rights (including intellectual property rights) may furnish background or default rules that govern the transaction unless otherwise agreed.”\(^{50}\)

A fourth argument is that the property rights of previous owners should trump the property rights of subsequent owners: “first owners determine the rights of second owners.”\(^{51}\) This is the ‘priority of rights’ argument. The fifth and final argument is that as long as all parties are fully informed of asset restrictions, there should be no legal barriers to their enforcement: “public regulation, either judicial or legislative, is to provide notice by recordation of the interest privately created.”\(^{52}\) This is the ‘notice cures all’ argument.

It is easy to see why notice-based enforcement of use restrictions would be desirable for sellers, particularly those of mass produced products: they need only affix a label to a product to make their desired restrictions enforceable against all downstream buyers. But, as we will see below, the law continues to refuse to enforce use restrictions based on labeling alone.\(^{53}\) Does the user innovation theory of the *numerus clausus* provide a justification for the law’s interference with such a desirable business strategy?

### iv) Restrictions on the Creation of Novel Servitudes Promote User Innovation

Although the restrictions on the creation of new servitudes, particularly on chattels, are difficult to understand in terms of freedom of contract, the priority of property rights or notice, they do make sense when viewed in terms of user innovation. The most direct form of user innovation occurs when users use assets in new ways, many of which have not been contemplated by manufacturers. Other new uses may have been contemplated by manufacturers, but been deemed undesirable by them. Thus, it is particularly important that the possessors of assets be able to use assets in any way they desire, an ability which has traditionally been included in the general right of full use or full user.\(^{54}\) The right to reuse an


\(^{51}\) Robinson, *supra* note 43 at 1462. See also Raymond Nimmer, *supra* note 50 at 854: “the essence of any property right is that the owner can transfer or withhold transfer of its property.”

\(^{52}\) Epstein, *Notice, supra* note 43 at 1354.

\(^{53}\) *Infra* Part 2, ‘Bare Notice’ Chattel Cases Remain Anomalous.

\(^{54}\) See e.g. E.C.S. Wade, *Restrictions on User*, 44 L. Q. REV. 51 (1928).
asset as many times as possible throughout its durable life is another simple aspect of the right to full use, but it plays an important role in fostering user innovation. This is because both problems and solutions are often revealed as a natural by-product of use. The right to reuse thus provides a trial and error process by which these problems and solutions can reveal themselves.\textsuperscript{55} Preserving the ability of users to use assets in new ways, and to repeatedly reuse assets, thus provides a countervailing benefit which justifies the interference with the pre-existing contract and property rights of the seller, even in the face of clear notice to the buyer.

User innovation also provides an answer to the criticisms that servitude law is archaic, complex and unfair. Servitude law is anything but archaic, because user innovation is \textit{more}, rather than less, important in our modern economy. User innovation also provides some justification for concerns about the complexity of servitude law. Behind the surface complexity of servitude law actually lies a unified goal – that of limiting the proliferation of restrictions on the use of assets. This goal clearly promotes user innovation.

The accusation of unfairness arises because servitude law can be circumvented in complex commercial transactions or inheritance arrangements - via highly sophisticated trust instruments or complex negotiated contracts.\textsuperscript{56} Although these mechanisms \textit{do} allow an end run around servitude law, they are costly to implement, and they are subject to their own internal limitations.\textsuperscript{57} As a result, individuals are allowed to circumvent servitude law in only a narrow set of circumstances, and where it is worth the time and expense for them to do so. This means that the vast majority of asset transactions remain subject to the legal restrictions on the creation of servitudes - and this fosters user innovation.\textsuperscript{58}

\textsuperscript{55} See discussion \textit{infra} Chapter 2, Part 2, \textit{Asset Use Provides a Trial and Error Process.}


\textsuperscript{57} The most obvious limitations on the use of contract, for example, are the requirements of agreement, privity and consideration.

\textsuperscript{58} For more on this channeling effect of the law, \textit{see infra} Part 2, \textit{User Innovation and the Mechanisms of Governance.}
2) The Law Has a Greater Tolerance for Servitudes on Land Than for Servitudes on Chattels

Although both servitudes on land and servitudes on chattels continue to receive judicial scrutiny, the law does not police these two types of assets with the same vigor. Instead, the common law’s concern with the creation of novel servitudes is particularly heightened with respect to personal property. This differential treatment of land and chattels was perhaps best expressed by the court in John D. Park v. Hartman: “The right of alienation is one of the essential incidents of a right of general property in movables, and restraints on alienation have been generally regarded as obnoxious to public policy, which is best served by great freedom of traffic in such things as pass from hand to hand… . A covenant which may be valid and run with land will not run with or attach itself to a mere chattel.”59 In other words, the law polices the numerus clausus on the basis of asset type. We have already seen that the common law strategy of limiting the number and variety of servitudes makes sense in terms of user innovation. But can user innovation also justify the fact that the law varies its approach to servitudes depending on the type of asset involved? As we shall, see, this apparent incongruity is fully explicable in terms of user innovation, because the need to search out new uses of assets varies by asset type.

While some judges have expressed doubt about the wisdom of the differential legal treatment of servitudes on chattels,60 the majority continue to support the policy.61 As is the case with the servitude issue more generally, scholars are also sharply divided on the issue. Some support the continuing legal treatment of servitudes on chattels,62 while others have been even more doubtful about the policy than the courts have been.63 It has even been suggested that the law should be less wary of servitudes on chattels, because most chattels do

60 See e.g. Tri-Continental, supra note 34 at 627: “Under the circumstances obtaining here, a great deal more should be put forward by plaintiff, in support of its claim, to rid the vessel and the parties dealing with it of the covenant, than the dry as dust and technical common law distinction between chattels and realty.”
61 See e.g. the dissent in Tri-Continental: “Believing, as I do, that there is a rational basis of continuing validity for the traditional difference between the principle of law which permits the burdening of real estate with restrictive covenants and that which favors the sale of tangible personal property unencumbered by even known restrictive covenants.” Id. at 626-7.
62 For much the same reasons that other scholars have supported the continuing restrictions on the creation of new land servitudes – i.e., notice and information costs, ‘the problem of the future’, and externalities. See Molly Van Houweling, The New Servitudes, 96 GEO. L.J. 885 (2008) [New Servitudes]. For reasons already explained, supra Chapter 2, Part 1, Departures from Previous Theories, the theory offered here neither adopts nor endorses any of these rationales.
63 See generally, Robinson, supra note 43.
not last as long as land, and this makes most chattel servitudes of relatively short duration. According to Robinson:

“One argument was that the severity of the restraint might be greater for personal property than for land because the former is transferred more frequently than the latter, which stays in the same hands for relatively long periods of time. It is an interesting argument, but quite backward. What counts is not product turnover but the durability of the restriction. A restriction on the use (or sale) of Blackacre can limit the use of a valuable resource for a very long time. Although deed covenants are subject to termination for changed conditions, and in some states by fixed statutory limitations, the duration of such covenants can be very long. Most restrictions on chattels will be self-liquidating in a short period of time simply because of the relatively shorter lifetime of the physical object.”

Another of Robinson’s objections is that, unlike a land servitude, a chattel servitude does not remove a unique good out of commerce: “More important than the point about duration is the actual effect. A personal property restriction rarely removes a unique product from commerce. If the owner of Blackacre imposes very restrictive conditions on the use of the land, a unique resource may be effectively removed from seeking its highest valued use. Nothing like that occurs in the case of chattels, at least those chattels that are the objects of servitude-like restrictions.”

When viewed in terms of duration and the removal of a unique good from commerce, the differential legal treatment of land and chattels does seem difficult to justify. However, when viewed through the lens of user innovation, it makes perfectly good sense. The key to the judicial discomfort with chattel servitudes is not the duration of the servitude, the number of times the asset is transferred, nor the removal of a unique good from commerce. Rather, it is the impact of the servitude on subsequent asset use (user innovation). The negative impact of a servitude on user innovation increases as the number of unique individuals subject to the servitude also increases. In the case of land, the ‘world’ may be theoretically subject to a servitude on a parcel of land, but there is generally only one possessor of the asset, so the impact on user innovation, in terms of its encumbrance of human capital, is relatively low.

---

64 Robinson, supra note 43 at 1489.
65 Id. at 1490.
Thus, the uniqueness of each land parcel actually makes land parcels relatively uninteresting from the point of view of user innovation.

With chattels, on the other hand, the situation is quite different. In fact, it is the non-uniqueness of chattels that creates a situation that is highly conducive to user innovation. The sale of many non-unique chattels allows the simultaneous use of an identical asset by different users. This generates different patterns of use, which leads to differential problem finding and problem solving across users.\(^{66}\) Thus, the potential for user innovation increases with the number of non-unique goods released onto the market. Of course, in the case of a truly unique chattel, such as original work of art that exists in only one copy, there is little reason for the law to interfere with restrictions on its use and disposition. But allowing a seller to implement a servitude scheme on a mass-market good would interfere with the process of user innovation, and this provides ample justification for the law to be more wary of servitudes on chattels than on land.

The sheer number of goods released into the market is an important determinant of the potential for user innovation, but it does not provide us with the full picture of the need to search out new uses of assets. The need (or lack thereof) for search can be divided into two broad categories: 1) search is desirable, and users have advantages in undertaking that search; and 2) search is undesirable. 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 assets that fall within this category. In the second 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 implemented via property customization are generally unobjectionable, and the *numerus clausus* should be at its weakest in this category. We will discuss each of these categories of search in turn.

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

Where there is a need to search out new uses of assets, there will be benefits to allowing the possessor of the asset to operate under a rule of exclusion with respect to those assets. In general, there is a substantial need to search out new uses of tangible, movable

\(^{66}\) *See generally Chapter 2, Part 2.*
assets: finding new uses of resources, making asset modifications, and creating interoperable products is critical to ongoing innovation in regards this type of asset.\textsuperscript{67} Thus, we see that the policy against servitudes on chattels coincides with the need for user innovation.

Although rules of exclusion also impose a short-term cost on those manufacturers who wish to enforce use restrictions and other governance rules, in the long run preserving the rights of users of assets benefits both users and manufacturers.\textsuperscript{68} In many cases, user innovations will be revealed and will be adopted by fellow user-innovators and even manufacturers. In these circumstances, 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. Where there is no incentive for a manufacturer to respond to a user’s needs, there is simply no good reason to enforce a use restriction.

In addition to there being a need for search in regards to these types of assets, there is also an advantage in allowing users to undertake the search for new uses of these assets. We have already examined (in Chapter 2) several case studies of user innovation, and have 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 those reasons here.\textsuperscript{69} The relative advantages that users have in problem finding and problem solving are greatest in regards to assets that are susceptible to repeated and highly variable patterns of use. Again, we find that many chattels fall squarely within this category of assets.\textsuperscript{70}

\textsuperscript{67} For evidence and data, see generally Chapter 2.
\textsuperscript{68} So why do some manufacturers seek to impose restrictions on users’ ability to innovate via the use of assets? For all the reasons set out in Chapter 3, Part 2: feelings of psychological ownership, the endowment effect of legal ownership, noneconomic motivations and shortsightedness. All of these phenomena cause manufacturers to focus on the short-term downside of user innovation (as a threat to their chosen business model), rather than the long-term upside (a potential source of new innovations). And there may be at least two other reasons that manufacturers seek to prevent user innovation: first, they may simply be unaware of the benefits of seeking out and adopting user innovations; and second, feelings akin to psychological ownership or an endowment effect may develop with respect to the business model itself, rather than just the assets that are manufactured and sold on the market.
\textsuperscript{69} These reasons are set out in Chapter 2.
\textsuperscript{70} Actually copies of intellectual property works also fall within this category, but the need to provide incentive for the production of the intangible work complicates matters. For a concrete example of how protecting users’ rights in copies facilitates innovation, see Chapter 5, \textit{The ‘Mere’ Right to Rent/Resell Fosters User Innovation}. 

148
ii) 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 principal 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.71 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.

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 the coordination of land uses; and 2) when we want to facilitate the expression of non-market values. The first category, the need to coordinate land uses, provides us with the basic rationale behind land servitudes, such as easements and covenants. The second category, the expression of non-market values, provides us with the motivating force for the creation of conservation easements, which restrict economic development in order to protect and preserve historic or environmental features of real property. We will discuss each of these in turn.

a) Coordinating Land Uses

Land servitudes open up the *numerus clausus* because they separate specified rights of use from the general right of possession. The former owner gets the specified use rights, while the possessor gets ‘everything else’. The specific parameters of these use rights vary greatly: they can be an actual right of use, such as a right of way which gives a county the

---

right to establish a highway that crosses a railroad track;72 or they can take the form of a right to prevent the possessor of land from making certain uses of his property, such as the ability to construct a building high enough to block an adjacent owner’s sunlight.73 The existence and variety of land servitudes provides clear evidence that the *numerus clausus* does not view land with the same paternal eye with which it views chattels. But this differential approach is in full accord with the user innovation theory of the *numerus clausus*.

The opening up of the *numerus clausus* to land servitudes facilitates the coordination of land uses. We can see this facilitation in both the types of servitudes allowed, and in the internal requirements for various types. For example, English courts recognize only four types of negative easements: those preserving the flow of water to an adjacent parcel, those allowing a flow of air from a defined channel, those allowing an easement of light, and those preventing the removal of lateral support for a building.74 The usefulness of these easements in coordinating land uses cannot be in doubt. Similarly, when we consider many of the internal requirements for the availability of real covenants and equitable servitudes, the goal of facilitating land uses is also apparent. For example, both the touch and concern requirement,75 and the requirement that there be a dominant and servient tenement,76 are animated by the goal of coordinating land uses. By allowing adjacent landowners to coordinate their activities, land servitudes facilitate efficient land development and foster a specialized type of user innovation - *coordinated* user innovation.

But the promotion of coordinated user innovation is only one of the benefits of land servitudes; the other is the prevention of *destructive* user innovation. For example, you would not want your adjacent suburban landowner to tear down his house and install a nuclear waste dump. If the zoning regulations of your community did not prevent this, then you and

---

72 See e.g. Central P. R. Co. v. Alameda County, 284 U.S. 463 (1932) (a county’s easement on a government corporation’s right of way).
73 See e.g. Bryan v. Grosse, 155 Cal. 132 (1909).
75 The touch and concern requirement for covenants was set out in Spencer’s Case, supra note 26. The role that the touch and concern requirement plays in facilitating the coordination of land uses and promoting the efficient use of land resources has been noted by several scholars. See e.g. French, Reweaving, supra note 12 at 1289; Reichman, Unified, supra note 28 at 1232-33; and Rose, Comments, supra note 15 at 1409-1410.
76 The requirement that an easement have a dominant and servient tenement can be found in Ellenborough Park, supra note 17.
your local landowners might want to implement such restrictions via a private servitude regime. By preventing negative externalities in the use of land, land servitudes both preserve the value of adjacent parcels and foster community harmony. Thus, the availability of land servitudes also helps to prevent and reduce negative externalities.

In sum then, there are three reasons why the ability of individuals to open up the *numerus clausus* in regards to real property does little harm to user innovation. First, the uniqueness of each land parcel means that land parcels are not subject to simultaneous use by many individuals in the same way that non-unique chattels are. As discussed above and in Chapter 2, this means that land does not generate opportunities for user innovation on the same scale as mass-market chattels. Thus, to the extent that land servitudes do hinder user innovation by restricting land uses, the deleterious impact of this is much more attenuated in the case of real property. Second, user innovation in the use of land parcels does occur, but the innovation is of a different type than that which occurs with chattels. As we have seen, the existence of land servitudes actually fosters this type of coordinated innovation. Third, the need to coordinate land uses means that, unlike other categories of assets, there is a downside to allowing individuals to freely innovate via use of their asset. Thus, land servitudes reduce negative externalities that can result from innovation in the use of land parcels. For all of these reasons, judicial recognition of land servitudes is fully in accord with the user innovation theory of the *numerus clausus*.

b) Expressing Non-Market Values

In addition to the common law servitude categories of easements, real covenants and equitable servitudes, there is a more specialized type of land use restriction, which is designed specifically to further the goals of historic preservation or environmental protection. Conservation and historic preservation easements are a more modern development than traditional land servitudes, created at both the federal and state level by statute. This

---

77 Supra, *The Law Has a Greater Tolerance for Servitudes on Land Than for Servitudes on Chattels*.
79 On the growing popularity of conservation easements, see e.g. Julia Mahoney, *Perpetual Restrictions on Land and The Problem of the Future*, 88 VA. L. REV. 739, 741-742 (2002). Despite their legislative acceptance, conservation easements are not without controversy. In particular, they have been subject to the charge that they
statutory implementation was necessary because conservation easements violate one of the most basic requirements of a restrictive covenant – that it could not be held in gross. That is to say, there is usually no land retained by the seller which is to be benefited by the conservation easement (i.e., there is no dominant tenement), and even if there was such a parcel of land, the restriction on development generally benefits the burdened land itself (i.e., the servient tenement).

Although land servitudes and conservation easements both involve efforts on the part of landowners to control the use of land after they part with its possession, these efforts generated very different institutional responses. The courts were willing to allow the creation of land servitudes (albeit with restrictions on the types and their manner of creation), but remained stubbornly resistant to the idea of conservation easements, thus making a legislative solution necessary. 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 question of time. Perhaps if given more time, judges might have eventually responded to changing social norms, and carved out new exceptions to the *numerus clausus* that were capable of handling the unusual characteristics of conservation easements.

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 and the concept of user innovation. Land servitudes promote user innovation that involves the coordination of land uses, and obstruct user innovation that would have harmful effects on neighboring parcels. Conservation easements, on the other hand, hinder land development in order to facilitate the non-market values of historic and environmental preservation. Unlike fail on their own terms because they are nothing more than clever ways to secure tax breaks for landowners. There have also been cases of more blatant commercial abuse, such as when golf courses employ conservation easements to ‘protect’ their decidedly profitable use of the land as an ongoing business concern. See the discussion in Nancy A. McLaughlin, *Conservation Easements - A Troubled Adolescence*, 26 J. LAND RESOURCES & ENVTL. L. 47 (2005). Despite these departures from pure altruism, conservation easements generally represent a landowner’s view that preservation of qualities such as historical value, natural beauty, wildlife habitat and species diversity is more important than economic development. This restriction on development is, of course, a restriction on the search for new uses of the real property in question. Furthermore, this restriction is significantly broader than most traditional servitudes, as it prevents a broad range of economic development of the burdened land, beyond specified residential or agricultural uses.

80 After a long period of resistance, easements held in gross are now frequently enforced in American law; covenants still generally are not. See THOMAS MERRILL & HENRY SMITH, PROPERTY: PRINCIPLES AND POLICIES (2007). It should be obvious that the term ‘conservation easement’ is a misnomer in both the legal and the functional sense of the word.

81 *Id.*
land servitudes therefore, historical and environmental preservation easements place the goals of land and community development below the goals of protecting the environment and preserving the past. 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. Thus, it is possible to understand the differential judicial response to land servitudes and conservation easements in terms of user innovation. Land servitudes, which promote coordinated user innovation and hinder destructive user innovation, were appropriate judicial additions to the numerus clausus; while conservation easements, which hinder land development, were not.

Now that we have seen how several features of the law of servitudes and of the policy against restraints on alienation directly promote user innovation, we will examine some facets of the numerus clausus that promote user innovation indirectly – by validating psychological ownership.

PART 2 VALIDATING PSYCHOLOGICAL OWNERSHIP

Clearly, the need for (and ability of) users to search out new uses of assets varies by asset type, and the law adjusts its level of interference with governance rules accordingly. Where there is little scope for user innovation, or where user innovation would have a deleterious impact on public safety, coordinated uses of real property or the production and disclosure of intangibles, the law provides more room for governance rules to flourish. Where assets are amenable to development by a variety of users, and the potential economic impact of user innovation is therefore high, rules of exclusion predominate.

However, asset type provides us with only one aspect of the law’s approach to determining the acceptability of governance rules or use restrictions: the other is provided by the characteristics of the transaction. The law employs two principal transaction characteristics to draw the line between governance and exclusion: the obligation of the possessor to return an asset to the owner of the use right; and the existence of a bargaining
process (and the nature of that process) between the possessor of the asset and the owner of the use right.

The use of these transaction characteristics to draw the line between governance and exclusion can be explained in terms of the validation of psychological ownership, because both 1) the permanence of possession, and 2) the nature of the bargaining process, play an important role in the development of feelings of psychological ownership. As we will see below, psychological ownership can account for the following three features of the *numerus clausus*: 1) the law’s aversion to divided rights but tolerance for fragmented rights; 2) the common law’s refusal to recognize permanent leases; and 3) the fact that notice does not determine the boundaries of the *numerus clausus*. We will discuss each of these features in turn.

1) **Divided Rights Versus Fragmented Rights**

In Chapter 3, we saw that anticommons theory treats divided rights and fragmented rights as equivalent, but that the law does not.82 Instead, the law treats divided rights (dividing use from possession) and fragmented rights (dividing assets across the tangible dimension or sharing ownership of a single asset) quite differently. The user innovation theory of the *numerus clausus*, by drawing on the phenomenon of psychological ownership, fully accounts for the law’s differential treatment of fragmented and divided rights. It makes sense for the law to treat divided rights and fragmented rights differently because feelings of psychological ownership operate quite differently in the case of fragmented and divided rights. We would expect that use and possession of half a widget would give rise to feelings of psychological ownership of just that – half a widget. Thus, psychological ownership would not interfere with a transaction requiring the combination of two halves of a widget.83

The situation is quite different when it involves a possessor of a full widget and an owner of use rights in that widget. In this scenario, both possession and legal ownership give rise to feelings of psychological ownership.84 The best way for the owner of the use right to

82 See Chapter 3, Part 4, *Psychological Ownership Versus Anticommons Theory*.
83 There may be other barriers to the formation of a deal in this scenario, of course. But these barriers involve disputes over value, and the *numerus clausus* does not concern itself with disputes over value, it only concerns itself with disputes over ownership. See Chapter 3, Part 4, *Value Versus Ownership*.
84 See generally, Chapter 3.
express those feelings may be via the denial of permission. The best way for the possessor of the asset to express those feelings is to use the asset without asking for permission at all. And since the possessor already has the full widget, in most cases he will simply be able to unilaterally violate the use right owner’s legal rights (this violation, in turn, creates a powerful incentive to hide the use). Together, all of these factors mean that barriers to recontracting are especially high when rights of use are divided from the right of possession. Unifying rights of use and possession completely eliminates some of these barriers to recontracting, and greatly reduces others.

2) The Common Law Does not Recognize Permanent Leases

When a court must classify a transaction as a lease or a sale, one of the most fundamental guidelines it uses is the fact there is no such thing as a ‘permanent’ lease: “At the expiration of a typical lease, possession reverts to the owner. A permanent lease, without reversion, is a permanent transfer of property rights, analogous to a fee simple transfer plus a permanent contract for services.” In addition to its importance in classifying leases versus sales, the nonexistence of permanent leases is used in eminent domain cases to determine whether there has been a taking: “that easement, if permanent and not merely temporary, normally would be the equivalent of a fee interest.” The fact that the law does not recognize permanent leases once again raises the question: why can a property owner not permanently transfer physical possession without parting with full ownership? Why must a lease always involve the return of the leased property? The answer, it will be shown below, lies in the impact that the duty to return an asset has on the development of feelings of psychological ownership.

86 United States v. Causby, 328 U.S. 256 (1946) at 262.
Possession and the Duty to Return

Possession has always had an important place in the law of property. *First possession* does much to determine initial ownership of unowned assets; and a long, uninterrupted period of possession can result in the legal change of ownership that we call adverse possession. But possession is pivotal in categorizing property transactions as well, and when possession of an asset has been permanently transferred, there is a strong legal presumption that such a transaction is a property transaction: “The right to perpetual possession is a critical incident of ownership.” Once a transaction is classified as a property transaction, the strong legal preference in favor of exclusion and against governance rules, also applies. When an asset must be returned to a use right owner, on the other hand, the law shows a much greater tolerance for the implementation of governance rules (and the consequent division of ownership).

One result of this approach is that there is no such thing in law as a permanent lease. Leases divide ownership between the individual who possesses the asset (the lessee), and the individual who retains various rights in the asset (the lessor). But one of the most defining characteristics of a lease is that the asset must eventually be returned to the lessor. If there is no duty to return the asset, then the law defines the transaction as a sale (and generally disallows use restrictions on the asset, or enforces them only against the original contracting parties).

The duty to return the asset has also played a defining role in characterizing a transaction as a license or a sale. The lack of a duty to return the copies of software programs, for example, played a key role in frustrating the attempts of software

---

87 See *e.g.* Robinson v. Campbell, 16 U.S. 212 (1818): “first possession without any other title, serves in an assise for land. In Bateman v. Allen, it was held that the plaintiff was entitled to recover in ejectment where it was found by special verdict that the defendant had not the first possession, nor entered under title, but upon the plaintiff's possession. And in Allen v. Rivington, where, upon a special verdict in ejectment, it appeared that the plaintiff had a priority of possession, and no title was found for the defendant, Saunders says, the matter in law was never argued, for the priority of possession alone gives a good title to the lessor of the plaintiff against the defendant, and all the world, excepting against the rightful owner.” (Citations removed.)

88 Incidentally, adverse possession is also consistent with psychological ownership.


90 So-called permanent leases on public housing have occasionally been created by statute, but they are generally permanent residential leases which are terminable for cause. Thus, such leases are not really permanent at all, and they are created for the special purpose of protecting low-income tenants from eviction. In most cases, even these leases have been repealed. See *e.g.* Otto J. Hetzel, *Asserted Federal Devolution of Public Housing Policy and Administration: Myth or Reality*, 3 WASH. U. J.L. & POL’Y 415, 432 (2000).

manufacturers to control the use of its software in ways that went beyond the U.S. Copyright Act.\textsuperscript{92} And even in those cases where the duty to return was hidden in standard form license terms but was never enforced, the judges looked to the ‘reality of the transaction’, and deemed the permanent transfers of copies to be sales, rather than licenses.\textsuperscript{93}

It has been suggested by scholars that a change in physical possession should not trump the pre-existing legal rights of owners who wish to permanently transfer physical possession of an asset without extinguishing all ownership rights in that asset.\textsuperscript{94} And indeed, when viewed in terms of the priority of rights, the legal prohibition on permanent leases seems capricious and unfair. However, when we consider the impact that possession and use has on psychological ownership, we can see that there is good reason for the law to continue to use the duty to return as a legal pivot point. In Chapter 3, we saw that possession facilitates all three of the functional pathways to the development of psychological ownership: controlling an asset, getting to know an asset and investing the self in an asset.\textsuperscript{95} Furthermore, possession enables use - which we saw was an even more important means of controlling, getting to know and investing the self in an asset. But, in addition to the mere fact of possession, the duration of possession is also important from the point of view of psychological ownership. The duration of possession is critical for at least two reasons. First, a longer term of possession simply provides more time to develop feelings of psychological ownership. Second, and just as importantly, individuals do in fact draw a psychological distinction between permanent and merely temporary possession. For example, we know that when we ‘rent’ a movie or a carpet cleaner it is not ‘ours’ because we must return it to its rightful owner.

Of course, the duty to return an asset does not always provide a completely satisfactory means to categorize a transaction. Extremely long lease terms on real property and commercial assets (terms of 99 or 999 years are well-known) can make the distinction

\textsuperscript{93} For more about the classification of a transaction as a license or a sale, see \textit{infra} Part 2, \textit{IP Copies and Notice: The Realities of the Transaction}.
\textsuperscript{94} See \textit{e.g.} Epstein, \textit{Notice}, \textit{supra} note 43; Robinson, \textit{supra} note 43; and Nimmer, \textit{Barriers}, \textit{supra} note 50.
\textsuperscript{95} Jon Pierce, Tatiana Kostova & Kurt Dirks, \textit{The State of Psychological Ownership: Integrating and Extending a Century of Research}, 7 REV. GEN. PSYCH. 84, 92-93 (2003). \textit{See also} the discussion in Chapter 3, Part 1, \textit{Psychological Ownership, Possession and Use}. 

157
between a lease and a sale seem like a game of legal formalist puffery. With terms approaching what we might call psychological levels of permanence, the fact that we cannot have a ‘permanent’ lease can seem like an artificial legal construct. However, most of these ‘long lease’ cases are residential in nature, and so they actually validate the feelings of psychological ownership of individuals who cannot or do not wish to purchase a dwelling (because they protect tenants from eviction). 

Furthermore, in those cases where individuals are subject to long leases - leaseholds in England, Wales and Scotland provide well-known examples - such arrangements are the locus of great controversy and ongoing attempts at legislative reform. Thus, even though long leases can protect tenants and validate the feelings of psychological ownership that arise from possession that is lengthy but not permanent, when such leases become so long that they mimic permanent possession, they generate demands for full ownership (‘conversion’ or ‘enfranchisement’).

3) Notice Does not Determine the Boundaries of the Numerus Clausus

Notice plays a pivotal role in previous economic theories of the numerus clausus, but in contrast to what these theories predict, effective notice does not necessarily result in the legal enforceability of use restrictions or other governance rules. In fact, the law not only frequently refuses to enforce restrictions when notice is relatively costless (a scenario that would justify the legal recognition of constructive notice), it often refuses to enforce restrictions even when there is actual notice. Below, we first examine how courts decide when notice will result in the enforcement of use restrictions. Then we examine how scholars have analyzed the issue of notice, and how they have criticized the law’s approach. Finally, we show how the user innovation theory of the numerus clausus successfully accounts for the legal approach to notice.

---


99 See e.g. Martin Davey, The Onward March of Leasehold Enfranchisement, 57 MODERN L. REV. 773 (1994).

100 For an explanation of why this is so, see infra, 3 B) The Scholarly View of Notice.
A) The Legal Approach to Notice

While it is true that recordation systems have resulted in some relaxation of the *numerus clausus*,\(^{101}\) the law continues to resist the idea that notice alone should determine the legal enforceability of use restrictions and other divided rights. Indeed, even when courts move from law to equity (where the presence of notice is most persuasive), notice alone is not usually sufficient to result in the enforcement of use restrictions. Instead, notice to subsequent transferees, *along with an enforceable contract between original covenantor and covenantee*, is usually required. We briefly look at whether courts appear to be loosening this contractual requirement below, in section i).\(^{102}\)

In section ii), we take a closer look at how notice affects the legal enforceability of use restrictions in the case of the sale of *copies* of intellectual property works.\(^{103}\) As we shall see, whether the asset in question is real property, a tangible chattel or a copy of an intellectual property work, courts do not merely look at the effectiveness of notice (that is, at whether the asset transferee has actual notice of use restrictions). Instead, courts examine both the nature of the transaction and the type of asset involved to determine if a use restriction is enforceable, and against whom (the original parties only, or all downstream transferees).

In the case of real property assets, the courts examine whether the restriction in question meets the requirements for the most closely related property form (whether that be an easement, covenant, servitude, or one of the types of permissible tenancies).\(^{104}\) In the case

---

\(^{101}\) See, e.g., Merrill & Smith, supra note 27 at 42.

\(^{102}\) See infra, ‘Bare Notice’ Chattel Cases Remain Anomalous.

\(^{103}\) Although most of our discussion of intellectual property (and of copies of such works) appears in Chapter 5, we discuss the ‘license or sale’ cases here because these cases require judges to decide between the rights of the possessor of the tangible copy, and the rights of the owner of the intangible property. This is the same task that judges are faced with in the chattel servitude cases (i.e., the intellectual property operates much like a dominant tenement, just like the business goodwill in the chattel cases).

\(^{104}\) Merrill & Smith provide a helpful example of how this procedure of property/contract classification works in practice: “[I]t is useful to consider an example of the principle in operation. Landlord-tenant law includes a version of the *numerus clausus* principle. Leases are limited to four recognized types: the term of years, the periodic tenancy, the tenancy at will, and the tenancy at sufferance. Suppose a landlord and tenant decide to enter into a lease that does not conform to any of the four standard types - a tenancy "for the duration of the war" being the classic example. If landlord-tenant law were just like the law of contract, then there would be no reason not to enforce this agreement in accordance with its terms; that is, the tenancy would last until the war ends. But courts typically do not proceed this way. Instead, they seek to determine which of the four recognized types of leases best fits what the parties have created. Since a term of years requires a "definite calendar
of tangible chattels, the courts sometimes determine the enforceability of restrictions under the rubric of equitable servitudes, while at other times they simply decide that the purported restriction in question is a contract rather than a sale (in which case the restriction is enforceable against the contracting party, but not against subsequent transferees). Similarly, the ticket/coupon cases, which involve non-transferability clauses, are sometimes analyzed in terms of equitable servitudes, and at other times are determined to be mere contracts.

However, many mass market movables have intellectual property content, and so the issue of post-transfer restrictions is often analyzed under the ‘license or sale’ rubric. To decide whether notice of a use restriction gives rise to a sale or merely a license, courts look to ‘the realities of the transaction.’ If the asset is merely licensed, there has been no transfer of ownership of the copy, and post-transfer restrictions can be enforced. On the other hand, if the asset has been sold, post-transfer restrictions cannot be enforced. To show how the ‘license or sale’ determination helps to police the boundaries of the numerus clausus, we discuss it in more detail in Section ii) IP Copies and Notice: The Realities of the Transaction.

i) ‘Bare Notice’ Chattel Cases Remain Anomalous

As noted in Part 1, scholars have raised several objections to the common law’s unaccommodating approach to servitudes, and two of these objections – the ‘priority of property rights’ and the ‘notice cures all’ arguments – do more than point to a more robust

---

105 As in Pratte v. Balatsos, supra note 33, where the dominant tenement was the business, and the servient tenement was the jukebox.
106 As in Sperry, supra note 37. In the ticket cases, the classification of the ticket/coupon as a contract generally results in its non-transferability, which results in the enforcement of the clause. This is in contrast with the tangible chattels cases, where the classification of a restriction as a contract also results in its non-transferability, but where this non-transferability means that the restriction is not enforced vis-à-vis downstream transferees.
107 Infra at ii) IP Copies and Notice: The Realities of the Transaction.
108 Insofar as the post-transfer restriction goes beyond the rights secured by the relevant intellectual property regime.
109 Infra Part 1, 1) iii) Scholarly Criticism of the Legal Policy Against Servitudes.
role for notice. They also support a fundamental reworking of the contractual foundations of equitable servitudes. In their original equitable incarnation, all servitudes have their genesis in a negotiated contract between two consenting parties. In contrast, the ‘notice cures all’ and the ‘priority of rights’ rationales for servitudes support the legal enforcement of servitudes that are grounded in mere labeling alone (i.e., neither the initial nor the downstream transferees of the asset would have to evidence contractual agreement to the use restrictions).

In both the originating case of *Tulk v. Moxhay*, and in later American cases, the equity of notice only attached to an asset where there was an enforceable agreement between the original covenantor and covenantee: “the question is... whether a party shall be permitted to use the land in a manner inconsistent with the contract entered into by his vendor, and with notice of which he purchased. Of course, the price would be affected by the covenant, and nothing could be more inequitable than that the original purchaser should be able to sell the property the next day for a greater price, in consideration of the assignee being allowed to escape from the liability which he had himself undertaken.”110 But if we take the arguments about the priority of property rights and the effectiveness of notice seriously, there is no reason for the law to continue to require an enforceable contract between the original covenantor and covenantee. After all, neither the priority of my property rights, nor the effectiveness of notice, depends in any way on the existence of a contract between the original transferor and transferee of an asset. Thus, arguments grounded in the priority of property rights or the effectiveness of notice support the enforceability of use restrictions based on the mere act of placing a label on an asset - no one need ever agree to the restriction contained in the notice.

Despite the demonstrated creativity of equity, the incidence of what might be called ‘bare notice’ (or mere labeling) cases remains rare. The case which created this legal form in the United States was *Waring v. WDAS Broadcasting Station*.111 Waring was the conductor of an orchestra that recorded some performances of musical works for Victor Talking Company. These records were sold with a label stating that they were: “Not licensed for radio broadcast.”112 Without this notice, Waring could not prevent WDAS from playing the

---

110 *Tulk v. Moxhay*, *supra* note 22 at 1144. The contractual moorings of equitable servitudes was affirmed in later English and American cases. Merrill & Smith, *supra* note 27 at 17.


112 *Id.* at 436.
orchestra’s recordings on the radio, because under the copyright laws sound recordings had no public performance right, and performers had no rights in their performances. Thus, WDAS only had to get a public performance license from the owner of the copyrights in the underlying musical works themselves.

The problem with this notice from the point of view of an equitable servitude on a chattel was that there was no contract between Waring and the Victor Talking Company restraining Victor from performing the works on the radio. The reason for this was simple: Victor Talking Company merely recorded the works and sold them in multiple copies – they were not in the business of performing works publicly. Of course, the recording company knew that Waring wanted to restrain the public performance of the records, and that is why they printed the notice on the label.113 Despite the absence of a corresponding contractual restriction between Waring and Victor Talking Company, the court enforced the radio broadcast restriction against WDAS radio station: “unless such a restriction can be imposed and enforced, it will be impossible for distinguished musicians to commit their renditions to phonograph records -- except possibly for a prohibitive financial compensation -- without subjecting themselves to the disadvantages and losses which they would inevitably suffer from the use of the records for broadcasting. Such a restriction, therefore, works for the encouragement of art and artists.”114

Waring remains something of an anomaly in the case law. However, its flexible approach to servitudes on chattels has been applied to a product that has little in common with vinyl records: hair color. In Clairol v. Sarann,115 the court built upon the foundation laid by Waring, enforcing a “Professional Use Only” label against a reseller who sold ‘professional use’ hair color to the retail consumer: “Waring is authority for the proposition that a purchaser with knowledge of a restriction is bound thereby, regardless of whether the restriction is contractual in nature--for in that case the restriction was enforced although it was not part of any contractual relationship between plaintiff and anyone in defendant's chain of title. This Court therefore holds that plaintiff is also entitled to the relief it seeks under the

113 “Defendant contends that there was no contract between plaintiff and the Victor Talking Machine Co. by which the latter agreed that the records should not be used for broadcasting purposes. There was, however, an understanding between them that the Talking Machine Co. would seek to prevent such use so far as lay within its power and would imprint the legend upon the records for that purpose.” Id. at 448.
114 Id. at 447.
theory of equitable servitudes." Although a few other courts have followed the Sarann case in enforcing such labeling restrictions against resellers (most for reasons that had nothing to do with equitable servitudes per se), their reliance on Waring is troubling. The Waring case was an unusual case of judicial creativity, motivated by a perceived gap in copyright law. But such issues should be addressed via reforms in intellectual property law: it is highly questionable whether judges have the information and expertise to create such IP-like regimes out of whole cloth. To date, Waring’s approach to chattel servitudes has only been adopted in a handful of cases. Given its unusual factual and legal background, courts should continue to be wary of adopting and extending Waring’s highly permissive approach to chattel servitudes.

Now that we have discussed the ‘bare notice’ cases, we will examine the ‘license or sale’ determination.

**ii) IP Copies and Notice: The Realities of the Transaction**

A paradigmatic example of the ‘license or sale’ determination is provided by the recent case of *UMG v. Augusto*. The recording company UMG created and distributed various promotional CDs and sent them out to music industry outsiders, in order to generate publicity in advance of the commercial release of new recordings. In order to prevent the promotional CDs from being recirculated, they printed the following language on the CD:

---

116 Id. at 735.
117 There’s no other way to say it: the judges that enforced servitudes on hair color did so because they were convinced by the plaintiff that home coloring her hair was an extremely dangerous activity for a woman. And perhaps in the 60’s and 70’s, it was. More modern courts have tended to allow the resale and home use of professional hair color as long as information sheets and warning labels are attached to the product by the reseller.
120 558 F. Supp. 2d 1055 (C.D. Cal. 2008) [UMG].

---
Many of the music industry insiders ignored the warning labels and sold their CDs, and Troy Augusto, who was not one of these insiders, purchased them from online auctions and music stores. He then resold them on eBay, advertising them as ‘rare collectibles not available in stores.’ This act of entrepreneurship came to the attention of UMG, and after first trying to halt the sale of the CDs through eBay’s VeRO program, which enables copyright owners to stop the auction of items that infringe their copyrights, UMG sued Troy Augusto for copyright infringement. Specifically, UMG asserted that Augusto had violated their exclusive right to “distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership.”

Augusto defended on the basis of the first sale doctrine, which states that notwithstanding the distribution right: “the owner of a particular copy or phonorecord lawfully made under this title… is entitled, without the authority of the copyright owner, to sell or otherwise dispose of the possession of that copy or phonorecord.” There was no question in the case that the promotional CDs were lawfully made; thus, the case turned on the question of whether Augusto was the lawful owner of the CDs. This in turn depended on the legal effect of the notice printed on the CDs. If the language was effective, then the CDs were merely licensed to the insiders, and the title to them remained with UMG. This would mean that the insiders could not lawfully transfer the CDs, and both Augusto’s acquisition of, and his subsequent resale of, the CDs would have been unlawful (i.e., he would not have been their lawful owner, and so his conduct would have been unprotected by the first sale doctrine).

---

121 Id. at 1058. Some of the CDs had a shorter version of this notice, but all of the labels purported to restrict redistribution of the CDs.
To determine the legal effect of the notice printed on the CD, the court took the time-honored approach of looking at the ‘economic realities’ of the transaction. That is, courts will not simply categorize a transaction as a license simply because such language is used in a contract or a printed notice. In addition, even though most of these notices use the term ‘license’, courts will not enforce the terms of a notice (whether it actually involves the term ‘license’ or not), even if the language of that notice has been read and understood by the transferee. Instead, they examine a variety of factors which serve to differentiate the permanent transfer of assets from the sort of ongoing obligations that characterizes licenses. The actual factors considered depend on the specific facts of each case, but the list of factors generally includes:

“In particular, the following factors require a finding… [that there has been a sale rather than a license]: temporally unlimited possession, absence of time limits on copy possession, pricing and payment schemes that are unitary not serial, licenses under which subsequent transfer is neither prohibited nor conditioned on obtaining the licensor's prior approval… and licenses under which the use restrictions principal purpose is to protect intangible copyrightable subject matter, and not to preserve property interests in individual program copies.”

In the *UMG* case itself, the court looked at primarily two factors: 1) whether the CDs had to be returned to UMG; and 2) whether the transaction created recurring benefits for UMG. Both of these factors supported the classification of the transaction as a sale or gift, rather than license: UMG mailed the CDs to the insiders with no expectation that they would be returned, and no continuing benefits flowed from the insiders to UMG (such as ongoing licensing fees, for example). The classification of the distribution of the promotional CDs as

---

124 *UMG*, supra note 120 at 1060: “In determining whether a transaction is a sale or a license, courts must analyze the "economic realities" of the transaction.” *See also* Microsoft v. DAK Industries, 66 F.3d 1091, n. 2 (1995): “Because we look to the economic realities of the agreement, the fact that the agreement labels itself a "license" and calls the payments "royalties," both terms that arguably imply periodic payment for the use rather than sale of technology, does not control our analysis.” *See also* SoftMan Prods. Co. v. Adobe Sys., 171 F. Supp. 2d 1075, 1084 (2001): “It is well-settled that in determining whether a transaction is a sale, a lease, or a license, courts look to the economic realities of the exchange.” [SoftMan].

125 SoftMan, *supra* note 124 at 1086. The court drew on the factors that were gathered from the cases by: David A. Rice, *Licensing the Use of Computer Program Copies and the Copyright Act First Sale Doctrine*, 30 JURIMETRICS J. 157, 172 (1990).
a gift rather than a license meant that UMG transferred title to the CDs when it mailed them to the industry insiders. Thus, despite the language printed on the CDs themselves, and despite the fact that no remuneration was received for the transfer, the language printed on the CDs was a legal nullity, and Augusto could not be prevented from reselling them. The UMG case provides an illustration of the fact that the *numerus clausus* does not police the proliferation of property forms unaided: much of the work of unifying rights of use and rights of possession is actually accomplished via the act of classifying a transaction as a license or a sale (or in the case of tickets/coupons and tangible chattels, via the act of classifying the transaction as a contract or property transaction).

Now that we have examined the common law’s view of notice, we will take a look at the view of law and economics scholars.

**B) The Scholarly View of Notice**

Because effective notice reduces the costs to parties within the chain of property transfer as well to third parties, 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 both property and contract law for a long time. As we have already seen, these debates have been particularly pointed in the literature on servitudes, with some commentators asserting that both the principles of private property ownership, and of freedom of contract, call for a regime in which effective notice is the *only* requirement for the creation of novel servitudes.126 These debates also play out in the intellectual property arena, with some scholars asserting that a mere printed notice that a copy is licensed rather than sold should be enough to trump the first sale doctrine.127

Notice also plays an important role in previous economic theories of the *numerus clausus*. Indeed, one of the primary differences between previous theories and the user innovation theory is that, in the latter, 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*,

---

126 See Epstein, Notice, *supra* note 43.
in which the cost of providing notice of unusual governance rules (property forms) plays a pivotal role in determining the need for the standardization of rights provided by 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. As Merrill and Smith put it: “As the costs of standardization… shift, we expect the optimal degree of standardization to rise or fall. Consider the rise of registers of interests in real property, that is, recording acts. This device lowers the costs of notice.”

Inherent 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 the case, then once notice cures the information asymmetry, by providing all parties with full information about the details of the governance rule, 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, they argue: “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.” (Emphasis added.) Information costs also play a key role in both Merrill and Smith’s, and in Hansmann and Kraakman’s, theory of the *numerus clausus*: “Merrill and Smith's contribution is important. We agree, in particular, that third-party information costs are central to the law's regulation of property rights.”

Clearly, law and economic scholars prescribe a more robust role for notice in determining the boundaries of the *numerus clausus* than that currently endorsed by the judiciary. Once again, however, the user innovation theory of the *numerus clausus* provides a justification for the current legal approach.

---

128 Merrill & Smith, *supra* note 27 at 40. See also Henry Hansmann & Reinier Kraakman, *Property, Contract, and Verification: The “Numerus Clausus” Problem and the Divisibility of Rights*, 31 J. Legal Studies S373, S374 (2002) [Hansmann & Kraakman]: “Property law generally… [presumes] that all property rights in a given asset are held by a single owner, subject to the exception that a partitioning of… adequate notice.”
130 Hansmann & Kraakman, *supra* note 128 at S373.
C) The User Innovation Theory Accounts for the Legal Approach to Notice

The user innovation theory of the *numerus clausus* accounts for the legal approach to notice in three ways. First, the user innovation theory shifts the focus away from information asymmetries (which shape the economic view of notice), and towards asymmetries in cognition and motivation (which implicitly shape the legal view of notice). Second, the user innovation theory’s focus on psychological ownership accounts for why contract (i.e., the presence of a bargaining process), but not mere notice, can defeat the *numerus clausus*. Third, the phenomenon of user innovation justifies the law’s strategy of allowing contractual enforcement of use restrictions even when it refuses to enforce *the very same restrictions* as a matter of property law, or on the basis of notice. We will discuss each of these issues in turn.

i) From Information Asymmetries to Asymmetries in Cognition and Motivation

If information asymmetries are the primary cause for concern in the *numerus clausus*, then the law and economics scholars are surely correct: effective notice should result in a relaxation of the doctrine. After all, notice will provide market participants with information about restrictions on assets, this information will result in changes in the relative prices of burdened and unburdened assets, and the market will work to rid us of any unwanted restrictions. But the argument here is that it is asymmetries in motivation and cognition, and not merely asymmetries in information, that are the real concerns that animate the *numerus clausus*.

There are two principal differences between information asymmetries and the kind of asymmetries central to the user innovation theory of the *numerus clausus*. First, information asymmetries are cured by notice, whereas cognitive and motivation asymmetries are not. Thus, there is no reason for the law to relax the *numerus clausus* simply because the cost of providing effective notice of use restrictions has decreased. Instead, in the user innovation theory, the *numerus clausus* expands in the face of notice only when user innovation is thereby fostered (or at the very least, when user innovation is not harmed, or when a noneconomic goal is at stake). Such is the case with land servitudes and conservation easements, respectively.131

131 *Supra* Part I, 2) ii) Search is Undesirable.
Second, each of these three types of asymmetries implicates a different type of problem (or problems). Information asymmetries present us with two problems: that of *measuring* what already exists, and that of communicating information that has already been produced. Cognitive asymmetries, in contrast, present us with the problem of *producing* what does not yet exist; they therefore impact innovative outcomes.\(^{132}\) As we saw in Chapter 2, asymmetries in cognition between users and sellers of assets, and between each individual user, mean that each individual user represents a unique source of potential innovation. Motivation asymmetries also have an impact on innovative outcomes, because a person’s *motivation to search* for new uses of assets will help determine whether there are any new uses worth striking deals over. Finally, the impossibility of predicting who will innovate (and how they will innovate) means that the optimal allocation of property rights cannot be determined *ex ante*.

Once user innovation has occurred, the information asymmetry view assumes that rights will be rearranged in response to revealed information about higher-value uses.\(^{133}\) *Ex post* to the innovative act, however, motivation asymmetries again come into play, because they impact subsequent bargaining behavior. An individual’s *motivation to deal* over new uses will determine whether value-enhancing deals are in fact struck (and whether such uses are revealed at all). As we saw in Chapter 3, the issues that arise under the heading of ‘motivation to deal’ are more complex than those that arise when information is simply costly to communicate or transmit to another party. Instead, the ‘motivation to deal’ category deals with the fact that psychological ownership, which arises both from physical possession and legal ownership, changes how individuals transact over assets. Because both the possessors of assets and the owners of use rights will not contract for rights which they feel that they already ‘own’, divided rights will frequently not be rearranged, even when such deals would make everyone better off.

\(^{132}\) It should be noted that, in this framework, cognitive asymmetries impact only information creation. That is to say, once information is produced and is made available via notice, it is assumed that there are no cognitive barriers to interpreting that information. This is not to suggest that these sorts of limitations do not exist; it is simply that they are of no interest in this theory. Such a cognitive barrier does appear in Merrill & Smith’s theory, for example (information is inherently costly to process).

\(^{133}\) These two assumptions essentially boil down to one: that innovative outcomes will be the same under different allocations of rights because individuals will rearrange rights to produce optimal innovative outcomes.
ii) Psychological Ownership, Notice and the Bargaining Process

In section 2 of Part 2, we saw that the term of possession (and the duty to return) plays a pivotal role in determining the boundaries of the *numerus clausus*. But the temporal nature of possession is not the only transaction characteristic that the common law uses to draw the line between governance and exclusion. The law also looks at the *process* by which governance rules are implemented – i.e., whether a bargaining process is involved - and this approach is fully consistent with the impact that bargaining has on feelings of psychological ownership.

Let us consider the easy case of ongoing, long-term contracting, where there is both interdependency and a repeated course of dealing. In such cases, governance rules are justified, not only because of actual contractual assent, but also because the ongoing awareness of legal rights, and the continuing compliance with legal obligations, is likely to interfere with the development of psychological feelings of ownership. Negotiated contracts, even if they represent discrete transactions, might also interfere with the development of psychological ownership, because they might act as *ex ante* ‘decoupling rituals.’ Recall from Chapter 3 that ‘decoupling rituals’ facilitate the termination of feelings of psychological ownership after they have developed. Although these rituals usually occur *ex post* to the development of feelings of psychological ownership, it is possible that the process of negotiation might prevent the development of psychological feelings of ownership. However, given the psychological power of subsequent use and possession, it is likely that only genuine, complex bargaining processes would be sufficient to prevent the development of feelings of psychological ownership.

At the opposite end of the spectrum is mere notice, such as a ‘do not use’ label on a chattel, or a ‘licensed not sold’ label on a CD. In these label cases, we have seen that users develop feelings of ownership toward such labeled assets, and that they expect to be able to

---

134 *Supra*, Part 2, 2) *Possession and the Duty to Return.*
135 The trucker example in Chapter 3 shows however, that even in ongoing relationships, feelings of psychological ownership can develop. See Chapter 3, Part 1, A) 1) *Psychological Ownership, Possession and Use.* However, it is arguable that in such cases, feelings of the extent of ownership are more nuanced. So, for example, even though the truckers felt sufficient ownership feelings towards their assigned trucks to maintain and personalize them, there is no evidence that they actually felt that the employer could not demand the return of the truck when their employment ceased.
136 Pierce et al, *supra* note 95 at 97. *See also* the discussion in Chapter 3, Part 1, *Psychological Ownership, Possession and Use.*
137 *Id.*
use them fully and freely. In other words, notice alone is not enough to interfere with the feelings of psychological ownership that develop via possession and use.

In between the two extremes of long-term contracting and mere notice are bargaining processes that fall short of an ongoing course of interdependency and repeat dealing, but go beyond passive notice. For example, if I must read and sign a contract at the time of sale, or if I must scroll through detailed terms and then click on an ‘I Agree’ button when I install software, these processes give rise to some level of awareness that full ownership of the asset has not transferred. However, these processes are not likely to counteract the development of feelings of psychological ownership if possession and use of the asset is uninterrupted.

Clearly, the spectrum of these intermediate cases is broad, and the contractual processes on this spectrum will interfere with the development of feelings of psychological ownership to varying degrees. Thus, it is to be expected that in these intermediate cases, contractual restrictions will sometimes be enforced even in the face of feelings of psychological ownership. In other words, psychological ownership dovetails quite nicely with the clear cut distinction between long-term contracting and mere notice, but provides a less reliable guide to the enforceability of use restrictions in the intermediate cases.

It is not surprising that cases intermediate between long-term contracting and mere notice are the most difficult for the user innovation theory of the *numerus clausus* to classify, because these cases are also the most troublesome for contract law. The tension between actual assent and mere notice has always strained the theoretical underpinnings of contract law, and for this reason the problem of standard form and adhesion contracts has occupied a substantial amount of scholarly and judicial attention. Indeed, the tension between assent and notice has increased rather than decreased over time, because technology continues to provide ever more cost effective ways to provide notice of use restrictions.

---

138 Recall the DivX and the DVD examples from Chapter 3, for example. See infra Part 1, A) 2) Individuals Expect Use to Travel with Possession.
140 See e.g. Michael J. Madison, *Legal-Ware: Contract and Copyright In The Digital Age*, 67 FORDHAM L. REV. 1025 (1998). We will explore these developments in Chapter 5.
iii) User Innovation and the Mechanisms of Governance

In section i), we saw that the user innovation theory of the *numerus clausus* focuses on asymmetries in cognition and motivation, rather than on information asymmetries. In section ii), we saw that psychological ownership accounts for the differential legal approach to contractual assent versus mere notice. In this section, we consider one final puzzle that surrounds the law’s approach to governance: the fact that the law does not regulate governance rules *per se*, but instead regulates the way in which sellers implement governance rules. That is, in the case of most use restrictions, the law does *not* allow or disallow the restriction based on the *content* of the restriction itself. Instead, the law regulates how the restriction is implemented. This has been the cause of recurrent criticism.\(^\text{141}\)

In our previous discussion, we saw that judges have usually refused to allow mere notice to give rise to either equitable or contractual obligations vis-à-vis chattels. On the property side, the policy against servitudes on chattels also thwarts private attempts to create new property forms. In the case of copies of intellectual property works, the courts look to the reality of the transaction to distinguish between licenses and sales.\(^\text{142}\) Of course, this does not mean that sellers can *never* limit the rights to resell or the right to full use: they are always free to contract with individual purchasers in order to restrict those rights. Indeed, the availability of the contractual mechanism of governance was indirectly pointed out in the *Quanta* case, which upheld the concept of patent exhaustion (the patent equivalent of first sale). The court was careful to note that exhaustion was only triggered in the case because “[n]othing in the License Agreement limited Intel’s ability to sell its products practicing the LGE patents.”\(^\text{143}\) Thus, the Supreme Court left the door open for sellers to restrict resale via licenses (contracts).

The fact that the law will often allow the enforcement of use restrictions via contract, but not via notice or property, has given rise to the criticism that the law is misguidedly

---

\(^{141}\) *Infra* note 144 and accompanying text.

\(^{142}\) In addition, courts have consistently read an inability to control the post-sale transfer of copies into the intellectual property grant itself. *See generally* Chapter 5, Part 1, 2) B) Copies are Like Chattels. Congress has also resisted attempts on the part of intellectual property owners to limit the first sale doctrine or the principle of exhaustion, except when there is clear evidence that the right to rent is being bootstrapped into a full-scale distributed piracy effort. Thus, the right to rent phonorecords of musical works and copies of computer software has been removed from the first sale doctrine. These ‘rentals’ were actually for the purpose of making reproductions. *Id.*

\(^{143}\) *Quanta Computer, Inc. v. LG Electronics, Inc.*, 128 S. Ct. 2109, 2122 (2008).
formalistic: “whether one should be able to contract around limitations on copyrighted or patented property should depend not on some formalistic distinction between contract rights and property rights, but on the policies at stake, and these policies require a closer examination.” Similar criticisms have been leveled at the first sale doctrine in intellectual property law: “[the first sale doctrine] is simply a limit on the owner's rights to claim patent or copyright infringement, leaving open the possibility that an owner might impose restraints by contract. The latter possibility seems to defeat the argument that the doctrine derives from alienation or trade policies, which should be no less applicable to enforcement of contract rights than to intellectual property rights.” Robinson’s criticisms, which have been echoed by other scholars, make sense: if a use restriction of a certain type is objectionable, why does the law not simply disallow that type of restriction? Why does the law permit a restriction via contract which it will not allow via property?

Behind this apparent incongruity lies a principled approach to use restrictions, one which makes sense in terms of user innovation. The common law does not regulate use restrictions or governance rules directly (except in the case of restrictions made illegal by other areas of law, and perhaps in the case of restrictions on resale). Instead, the law regulates the manner by which use restrictions are implemented and enforced – that is, the law regulates the mechanisms of governance. There are four ‘mechanisms of governance’ (i.e., mechanisms that manufacturers and sellers can use to implement governance rules): contract, notice, property and technology. The first three mechanisms of governance are constructs of law, and they are discussed together in this section. The final mechanism of governance is a creature of technology, and it therefore poses special challenges for legal regulation. We will discuss the technological mechanism of governance in Chapter 5.

Each of the legal mechanisms of governance has its own advantages and disadvantages for sellers, and each has received varied level of acceptance by the law. For example, the law is relatively tolerant of governance rules implemented via contract, but this

---

144 Robinson, supra note 43 at 1453.
145 Id.
147 Illegal contracts are one area where courts regulate the governance rule directly. See e.g. Sender v. Simon, 84 F.3d 1299, 1307 (1996): ‘courts generally will not enforce an illegal contract based upon “the elementary principle that one who has himself participated in a violation of law cannot be permitted to assert in a court of justice any right founded upon or growing out of the illegal transaction.”’ (Citations removed.)
mechanism is risky for sellers, because it requires manifestation of contractual assent and binds only the parties to the contract. In contrast, the law is highly intolerant of the implementation of governance rules via notice, even though this mechanism is highly desirable for sellers because it can be affixed to an asset relatively cheaply, and can easily bind anyone who sees it.

Of course, property is the most desirable mechanism of governance for sellers, because a property right requires no agreement for its enforcement, is good against the world, and travels with an asset. We are perhaps unused to thinking about property as a ‘mechanism of governance’, but that is only because we are accustomed to thinking about property in its numerus clausus incarnation. If property owners had complete freedom to customize novel forms of property (or to create new divisions of rights), then ‘property’ would actually be the ultimate legal mechanism of governance. Indeed, it is accurate to say that a fundamental goal of property law is to make sure that property only becomes a mechanism of governance (at least after an asset has been permanently transferred) in the most exceptional of circumstances.

From the point of view of sellers, the law’s receptivity to the mechanisms of governance is backwards. Contract is the most (legally) successful mechanism of governance, but it is the least desirable mechanism for sellers, because it requires that each buyer agree to the restrictions on resale or use; exactly the opposite is true for property. Considered in light user innovation however, this legal pigeonholing makes sense. By forcing sellers and manufacturers to rely on the leakiest and most costly mechanism of governance, the law strictly limits the number of assets in the marketplace that are subject to use restrictions. This greatly enhances the potential for user innovation to occur via the unrestricted use and resale of those assets. In addition, the common law strategy of funneling most use restrictions into the regime of contract provides some legal flexibility for those sellers who are willing to incur extra time and cost to implement use restrictions, while still maintaining the oversight of the property system.
CONCLUSION

In this Chapter, our detailed examination of the various legal doctrines that implement the *numerus clausus* has revealed that the law draws the line between governance and exclusion along two major axes: it limits the availability of governance regimes for certain *types of assets* (such as chattels), and for certain *types of transactions* (such as those involving permanent transfer of a physical object). These methods of differentiating between enforceable and unenforceable use restrictions have been criticized as artificial and arbitrary by some scholars (and a few judges). However, by viewing these methods of differentiation through the lenses of user innovation and psychological ownership, we can see that the legal approach makes sense. Adjusting the availability of use restrictions based on asset type promotes user innovation; while adjusting their availability based on transaction characteristics both promotes user innovation and validates psychological ownership.

Now that we have seen that the user innovation theory of the *numerus clausus* explains how the law regulates the availability of use restrictions on tangible assets, we will examine whether it can successfully account for one of the richest sources of divided rights: intellectual property law. We will also examine the future of the *numerus clausus* more generally. In particular, we will look at how the law should respond to technological developments that decrease the cost of providing notice of unusual use restrictions.
Chapter 5: The User Innovation Theory, Intellectual Property, and Technology

“In connection with the ambiguities latent in the term ‘property’, it seems well to observe that similar looseness of thought and expression lurks in the supposed (but false) contrast between ‘corporeal’ and ‘incorporeal’ property.”

1 Wesley Newcomb Hohfeld, Some Fundamental Legal Conceptions as Applied in Judicial Reasoning, 23 YALE L. J. 16, 23 (1913)
INTRODUCTION

In Part 1 of this Chapter, we examine whether the user innovation theory of the numerus clausus can account for the structure of intellectual property law, which is the most fertile source of divided rights or use restrictions on assets. In particular, we show that the user innovation theory can account for the following features of intellectual property law: the idea/expression dichotomy, the nonobviousness criterion, and the first sale/exhaustion doctrine. Then, in Part 2 we move on to discuss the implications of the user innovation theory for the future contours of the numerus clausus. In particular, we discuss whether advances in technology justify the relaxation or eradication of the law’s historical preference for exclusion, as previous theories of the numerus clausus have suggested. As we shall see, the user innovation theory of the numerus clausus leads not only to a justification of (and an explanation of) the law’s numerus clausus doctrines, but also to a call to preserve those doctrines in the face of technological developments. To the extent that the numerus clausus continues to face pressure to give way to privately crafted governance rules, the user innovation theory of the numerus clausus thus prescribes a steady course for legal doctrine, and provides a reasoned justification for that course. This in itself is a valuable contribution to legal doctrine and theory.

PART 1    THE USER INNOVATION THEORY AND INTELLECTUAL PROPERTY LAW

Because they interfere with the use and modification of assets, contractual use restrictions and servitudes are clearly antithetical to user innovation. However, intellectual property rights also allow sellers to prevent the use and modification of assets, via the public performance and the derivative work rights, for example. So does this mean that the user innovation theory of the numerus clausus implies that intellectual property law is also ill advised?

2 The public performance right is set out in 17 U.S.C. § 106(4): “the owner of copyright under this title has the exclusive rights to do and to authorize any of the following… in the case of literary, musical, dramatic, and choreographic works, pantomimes, and motion pictures and other audiovisual works, to perform the copyrighted work publicly.” The derivative work right is set out id. at § 106(2): “to prepare derivative works based upon the copyrighted work.”
Uncertainty about whether intellectual property rights promote or hinder the creation and dissemination of artistic and inventive works has always permeated the debates about intellectual property law. This uncertainty has never been adequately resolved by either the theoretical or empirical research, and it surely cannot be resolved here. Instead, we will assume that intellectual property law promotes some level of innovation that is socially beneficial, and examine what the user innovation theory of the *numerus clausus* has to say about the parameters of those rights.

In Chapter 4, we saw that the *numerus clausus* does not treat all assets equally; instead the law varies its tolerance for use restrictions based on the type of asset involved. We have already seen that chattels and land exemplify two categories of asset type, and these two asset types warrant a differential legal tolerance for use restrictions because they facilitate different kinds and intensities of innovation in the search for new uses of assets. In the case of chattels, a large number of non-unique assets, along with differential patterns of use and innovation across users, create a situation in which users generate a significant amount of innovation. We therefore referred to this category of search as: *Search is Desirable and Users Have an Advantage in Undertaking that Search.* In the case of land, the uniqueness of land parcels, the need for coordinated innovation and the need to prevent negative externalities, mean that the need to search out new uses is diminished. We thus referred to this category of search as: *Search is Undesirable.*

Intellectual property assets provide us with a third type of asset, and a third category of search. In this third category, search is desirable, but special asset characteristics dictate that non-users should control the search for new uses. We will therefore refer to this category as: *Search is Desirable, but Non-Users Should Control the Search.* The need for non-users to control or coordinate the search for new uses warrants some loosening of the *numerus clausus*, because it requires that non-users retain use rights in assets that have been physically transferred to users. Thus, we would expect *some* legal sanction of governance rules in this third category. As we shall see, because the assets in this category have some of the

---

4 See Chapter 4, Part 1, *The Law has a Greater Tolerance for Servitudes on Land Than for Servitudes on Chattels*.
5 Id.
6 Id.
characteristics of land and some of the characteristics of chattels, it is the most difficult category to respond to legally, and this category generates the most controversies in courts, legislatures and markets.

This Part is divided into two sections. In the first section, Search is Desirable, but Non-Users Should Control the Search, we briefly discuss the new category of search that is created by the special characteristics of intellectual property assets. In the second section, The User Innovation Theory and the Structure of Intellectual Property Law, we examine how the user innovation theory of the numerus clausus accounts for the current structure of intellectual property law.

1) Search is Desirable, but Non-Users Should Control the Search

In general, the user of an asset will have advantages in undertaking a search for new uses of, or modifications to, that asset - whether or not that asset has intellectual property embodied within it. In other words, intellectual property owners do not necessarily have advantages in undertaking the search for new uses of their intangible assets. Kevin J. Anderson might, for example, be able to write a better Harry Potter sequel than J.K. Rowling. However, in the case of intellectual property assets, we limit the ability of users to exploit their natural advantages in order to encourage the initial production and disclosure of intellectual property works. Intellectual property law allows intellectual property owners to control the search for new uses (for example, the ability to create, or to authorize the creation of, derivative works), in order to elicit the initial production of a nonrival asset that can be duplicated at a very low cost once the first copy of it is produced. If we do not allow the creator of the initial work to control (and profit from) subsequent reproduction, use and

---

7 See generally, infra Chapter 2.
8 But thankfully, we’ll never know.
9 In other words, the marginal cost of producing the initial copy of the work is very high, whereas the marginal cost of producing additional copies is close to zero. This means that without intellectual property rights, the creator of an initial work risks losing his entire investment once the work is disseminated to the public. See e.g. WILLIAM LANDES & RICHARD POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 40 (Harvard University Press 2003) [LANDES & POSNER].
adaptation of the work, the work may not be produced at all - this is the basic incentive rationale for intellectual property rights.  

It is important to provide incentives for the initial production and disclosure of intellectual property assets; however, we always need to keep a close eye on the scope of intellectual property rights. The nonrival nature of intellectual property works makes them ideally suited to being developed by many individuals. And so, as intellectual property rights get broader, the burden on user innovation increases. Traditionally, the scope of intellectual property rights has been controlled in two main ways: 1) by limiting the availability of protection via eligibility requirements; and 2) by limiting the post-transfer control of copies of works via various doctrines and defenses. For copyrighted works, this meant: 1) refusing to protect ideas and functional aspects of a work; and 2) giving the owners of copies some freedom to use and manipulate copies, primarily via the first sale doctrine. For patented inventions, this meant: 1) requiring that an invention be not just novel, but nonobvious; and 2) giving users some freedom to use and dispose of embodiments, primarily via the patent exhaustion doctrine. As we shall see in the next section, these methods of controlling the scope of intellectual property rights can be explained in terms of the user innovation theory of the *numerus clausus*.  

---

10 In addition to the incentive rationale, we may decide to allow authors and inventors to control the development of their works and inventions in order to recognize the personality interests of creators in their creations. See e.g. David Troutt, *I Own Therefore I Am: Copyright, Personality, and Soul Music in the Digital Commons*, 20 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 373 (2010). For the purposes of the user innovation theory of the *numerus clausus*, it does not matter why we protect intellectual property. It matters only that protecting intellectual property hinders user innovation of a nonrival resource.

11 Many works make this point. See e.g. Mark Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 998 (1997): “[T]he existence of preexisting intellectual property rights imposes a positive cost on improvers that they would not otherwise face. It is not enough to say that intellectual property law favors "creators" - for here we have creators on both sides of the equation, and the law must choose between them.... [T]he intellectual property rules governing improvements are important in understanding the extent to which protection for first-generation innovation will impede improvement in subsequent generations.” (citations removed).


13 There is a third major type of limitation – the limitation on the *duration* of the intellectual property right. However, because we are concerned only with limitations that apportion rights as between owners and users during the term of intellectual property protection, duration will not be discussed. Other scholars have said that duration should be the only limitation on an intellectual property right: see Richard Epstein, *The Disintegration of Intellectual Property? A Classical Liberal Response to a Premature Obituary*, 62 STAN. L. REV. 455, 509 (2010).

14 There are other ways in which intellectual property law limits the scope of intellectual property rights - the fair use doctrine in copyright law, for example. 17 U.S.C. § 107. However, these other limitations on intellectual property rights will not be discussed in detail in this thesis (although they are mentioned) because,
2) The User Innovation Theory and the Structure of Intellectual Property Law

From the point of view of the user innovation theory of the *numerus clausus*, controlling the scope of intellectual property rights via eligibility requirements and the first sale/exhaustion doctrine makes complete sense. In fact, these two legal strategies enable intellectual property law to regulate intangible assets in much the same way that property law regulates tangible assets. Recall that the law regulates the availability of governance rules based (in part) on asset type, with land and chattels being the two major types.\(^{15}\) Refusing to protect ideas and functional aspects of a copyrighted work makes the *work* more like a parcel of land; the nonobviousness criterion does the same in patent law. On the other side of the asset type equation, the first sale doctrine in copyright law (and the doctrine of exhaustion in patent law) makes the *copy* of the work more like a chattel. We discuss each of these assertions in more detail below.

A) Intellectual Property Works are Like Land

At first glance, it may seem that intangibles and real property have little in common with each other. However, an intellectual property *work* is akin to a parcel of land in three ways: 1) both give rise to a need for coordinated innovation; 2) both give rise to a need to control negative externalities; and 3) both are *unique* assets. The first two commonalities – the need for coordinated innovation and the need to control negative externalities - are the very *raisons d’être* of intellectual property rights. We have already discussed the basic incentive rationale for intellectual property rights: an asset which is costly to produce but cheap to reproduce might not be created at all unless we allow the creator of the initial work to control subsequent use, reproduction and adaptation of the work. But the basic incentive rationale only addresses the question of *who* will receive the monies from reproduction, use and adaptation. The need for *coordinated* innovation and the need to control negative

\[^{15}\text{Supra note 4.}\]
externally are somewhat more subtle concerns: an intellectual property owner needs to control exploitation of a work, not simply so that he will recoup sufficient monies from his initial investment, but also because he needs to prevent the overexploitation of a work. This phenomenon is often referred to in the literature as ‘congestion externalities.’

Congestion externalities - the need to coordinate innovation and to control negative externalities - arise from the nonrival nature of intellectual property assets. Because an almost unlimited number of copies of a work can be produced after its creation, and because an almost unlimited number of individuals can create adaptations of a work, two dangers arise. First, copies and adaptations of a work might flood the market, and the public might literally get sick of the work (and its adaptations) altogether. Second, some of these adaptations might be of poor quality, or they might be fundamentally inconsistent with the parent work. An adaptation is fundamentally inconsistent with a parent work when it introduces new elements that destroy the meaning or cohesiveness of a work. For example, if an unauthorized movie adaptation of the novel ‘Dune’ contained a Baron Vladimir Harkonnen who was selfless and moral, readers of the novel would surely perceive that something was amiss with such an adaptation.

Of course, inconsistent (and poor quality) adaptations can also be produced by authorized licensees or by the original authors themselves, but the likelihood of incoherence and incompatibility increases sharply if literally ‘the whole world’ can develop and exploit a work. Poor quality or inconsistent adaptations obviously risk their own failure in the market, but they can also harm the entire body of related works, because related works are associated with each other in the minds of consumers: “The result could be premature saturation of the market, consumer confusion (for example, as to the source of the derivative works), and impaired demand for the original work because of the poor quality of some of the unauthorized derivative works.”

---

16 See e.g Landes & Posner, supra note 9 at 226.
17 One infamous example is the ‘Felicity’ television series, where the producers thought that cutting the trademark long, curly hair of the lead actress, as part of a story line about a bad breakup, would be a good idea. This simple (yet profound) change in the character is widely believed to have caused the show’s plummeting ratings, which ultimately led to its cancellation. After the debacle, WB network executive Susanne Daniels vowed: “Nobody is cutting their hair again on our network.” See APPROVES: WEBSTER’S QUOTATIONS, FACTS AND PHRASES 406 (ICON Group International 2008).
18 Landes & Posner, supra note 9 at 226.
Not surprisingly, the need to coordinate innovation and control externalities does not operate in precisely the same way in intellectual property and in land. In the case of land parcels, coordinating innovation and controlling externalities occurs across *different land parcels*; whereas in intellectual property coordination and control occurs *within a single work* (and its related properties). However, from the point of view of the user innovation theory, this difference does not matter. The relevant point is that there is a need to foster coordinated innovation, and to prevent harmful innovation, and this leads to an opening up of the *numerus clausus* in both cases.

The third and final point of commonality between intellectual property works and land parcels – that both are unique assets - is not inherent to the intangible nature of intellectual property, but is instead a deliberate result of the legal structure of the rights. That is to say, intellectual property laws are structured so that works that receive intellectual property protection are unique. Copyright law ensures that the vast majority of copyrighted works will be unique in two ways: by protecting only the expression in a work but not the ideas;\(^\text{19}\) and by refusing to protect the functional elements of a work.\(^\text{20}\) Protecting expression but not ideas ensures that copyrighted works will be unique because it is highly unlikely that two individuals will express even the same idea in exactly the same way. In other words, expression will tend to *diverge*.\(^\text{21}\) On the other hand, functional elements by their very nature will tend to *converge*.\(^\text{22}\) Thus, protecting expression, but not ideas or functional elements, helps to prevent a situation in which granting protection to one work will potentially interfere with the development of another work. While it is true that copyright law does allow in *theory* for the protection of two identical works, in practice this will rarely occur. In addition, independent creation is a complete defense to copyright infringement.\(^\text{23}\) So, even in the case where identical works *are* created by different individuals, copyright law interferes with

---

\(^{19}\) *See e.g.* Feist Publ'ns, Inc. *v.* Rural Tel. Serv. Co., 499 U.S. 340, 349-350 (1991): “copyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work.”


\(^{22}\) *Id.*

\(^{23}\) Actually, this is not quite accurate. Independent creation means that no copyright infringement has occurred at all, so the plaintiff has not stated a *prima facie* claim of infringement. *See e.g.* Alfred Bell & Co. *v.* Catalda Fine Arts, Inc., 191 F.2d 99, 103 (2d Cir. 1951).
copying but not innovation - even where that innovation is merely the independent creation of an identical work.

Of course, functional elements and ideas can be protected by another body of intellectual property law - patent law. But patent law has its own set of techniques for ensuring that protected inventions are unique assets. In fact, patent law goes even further than copyright law in ensuring that only unique works receive protection. First, unlike copyright law, patent law refuses to protect identical inventions, even if they have been independently created.\(^{24}\) Second, patent law has a special doctrine, the doctrine of equivalents, which ensures that the protection for a patented invention goes beyond the literal claims of an invention, and also covers its functional equivalents.\(^ {25}\) This also helps to ensure that patented inventions are unique, because no one else can secure a patent for an equivalent invention. Finally, patent law uses the nonobviousness criterion to ensure that protected inventions are not just new and useful, but that they also represent an advance over prior inventions which would not be obvious to individuals of ordinary skill and knowledge within the field of the invention.\(^ {26}\) In other words, the nonobviousness criterion seeks to ensure that inventions that receive patent protection are inventions that we would not get from just any person skilled in the relevant art.\(^ {27}\) Thus, nonobviousness reduces the likelihood that simultaneous inventions

---

\(^{24}\) Patents are granted on either a first to invent basis (U.S.), or a first to file basis (most other countries). See e.g. Duncan Matthews, Globalizing Intellectual Property Rights: The TRIPS Agreement 58 (Routledge 2002).

\(^{25}\) The doctrine of equivalents was created by the courts in Winans v. Denmead, 56 U.S. (15 How.) 330, 334 (1854). Copyright law also protects against more than literal infringements of a work: the test for infringement is substantial similarity, not complete identity. This prevents copyists from escaping liability by making (qualitatively or quantitatively) insignificant changes to a work. See Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d. Cir. 1930). However, because copyright law also allows the independent creation of works that are not only substantially similar, but even identical, substantial similarity does not have quite the same bite that the doctrine of equivalents has in patent law.

\(^{26}\) 35 U.S.C. § 103(a).

\(^{27}\) This is the goal, but whether patent law succeeds in that goal is another question. Practical administration of the patent system seems to have resulted in a glut of patents that are obvious even to people not skilled in any art. See e.g. Justin Pats, Preventing the Issuance of "Bad" Patents: How The PTO Can Supplement Its Practices and Procedures to Assure Quality, 48 IDEA 409, (2008): “Another so-called "bad" patent comprised claims for a method of swinging on a swing. Invented by a five-year-old child whose father was a patent attorney, this invention should have been found trivially obvious. The patent recently underwent re-examination and was eventually cancelled as a result. Additionally, there was the "one-click" patent granted to Amazon for an e-commerce method of purchasing an item online with a single click of the mouse after the purchaser had previously entered necessary credit card information to complete the transaction. Many public observers viewed the patent as non-novel or obvious, especially since many companies had already implemented similar systems at the time of Amazon's filing. Amazon would eventually settle a lawsuit over the patent with competitor Barnes & Noble in 2002.” (Citations removed.)
will be *patentable* inventions\(^{28}\) – i.e., it helps to ensure that patented inventions are ‘unique’ in a way that goes beyond being merely ‘new’.

From our perspective, protecting only unique works and inventions is important because it helps to mitigate the negative impact that intellectual property rights have on user innovation. The very fact that copies of works can be distributed among a large number of users means that copyright law sacrifices the input of a large number of users - each of whom could potentially create derivative works. But by protecting only expression, which will tend to diverge across individuals, copyright law at least gives all individuals the chance to create and receive protection for *their own* original works. In trying to ensure the uniqueness of works, patent law has a different set of concerns, because the functionality of inventions will tend to make multiple innovators converge to solutions that are similar, if not identical. Patent law tries to at least partially offset the potential negative impact of granting a patent in such situations by refusing to protect innovations unless they have reached a certain quantum of inventiveness.

Of course, eligibility requirements are not the only techniques used by copyright and patent law to protect user innovation while also granting intellectual property rights. The fair use doctrine, for example, carves out a special place for user innovation in regards to copyrighted works by giving users the freedom to create certain kinds of derivative works, such as parodies.\(^{29}\) And patent law has an even more robust protection for user innovation, in the form of ‘blocking patents.’\(^{30}\)

Of course, in the case of intellectual property assets, the ability to reproduce a work means that although an intellectual property *work* has much in common with a parcel of land,

\(^{28}\) Nonobviousness does not eliminate the phenomenon of simultaneous patentable invention however, even when the patent examination procedure works as it should. For stories of simultaneous invention and races to the patent office *see e.g.*, HAL VARIAN, JOSEPH FARRELL & CARL SHAPIRO, THE ECONOMICS OF INFORMATION TECHNOLOGY 6, 68 (Cambridge University Press 2004).

\(^{29}\) *See e.g.* Campbell v. Acuff-Rose Music, 510 U.S. 569, 592 (1994): “there is no protectible derivative market for criticism.”

\(^{30}\) Suppose an innovator builds upon a patented invention in a manner that is itself patentable, but such that the improvement can only be practiced together with the underlying patented invention. A blocking patent gives the improver rights in his improvement, but neither the original patentee nor the improver can practice the improvement unless they license each others’ patents. In other words, the two patents ‘block’ each other. *See e.g.* Standard Oil Co. v. United States, 283 U.S. 163, 171 n. 5 (1931): “This is often the case where patents covering improvements of a basic process, owned by one manufacturer, are granted to another. A patent may be rendered quite useless, or "blocked," by another unexpired patent which covers a vitally related feature of the manufacturing process. Unless some agreement can be reached, the parties are hampered and exposed to litigation. And, frequently, the cost of litigation to a patentee is greater than the value of a patent for a minor improvement.”
the copies of that work are the very antithesis of real property. Thus, we also need to examine how intellectual property mediates the tension between user innovation and the intellectual property owner’s control of copies.

B) Copies are Like Chattels

One of the most fundamental distinctions in intellectual property law is between the intangible intellectual property, and the tangible embodiments of that intellectual property. Copyright law makes a fundamental distinction between the protected work (e.g., the literary work) and the copies of that work (e.g., a book). Similarly, patent law distinguishes between the invention, and the embodiments of that invention. This distinction makes it necessary to address the apportioning of rights as between the owner of the intellectual property work, and the owner of a copy or embodiment of that work.

We have already seen that the owner of the intellectual property right gets most (but not all) of the rights in the intellectual property work. As we shall see below, the owner of the copy gets many (but not all) of the rights in the copy. This apportioning of rights in

\[\text{\footnotesize[31 See 17 U.S.C. § 202: “Ownership of a copyright, or of any of the exclusive rights under a copyright, is distinct from ownership of any material object in which the work is embodied. Transfer of ownership of any material object, including the copy or phonorecord in which the work is first fixed, does not of itself convey any rights in the copyrighted work embodied in the object; nor, in the absence of an agreement, does transfer of ownership of a copyright or of any exclusive rights under a copyright convey property rights in any material object.”}\]

\[\text{\footnotesize[32 The copyright owner gets most, but not all rights in the work. Fair use, for example, has the effect of giving owners, possessors and even mere users of a copy the right to prepare certain kinds of derivative works, and the right to make certain copies. The fair use doctrine is codified in 17 U.S.C. § 107.}\]

\[\text{\footnotesize[33 Throughout this discussion, I generally refer to ‘the owner of the copy’ rather than ‘the possessor of the copy’ or ‘the user of the copy’ because an individual can violate the rights in a copy without violating the rights in the copyrighted work. For example, if I steal the book that you lawfully purchased, I violate your rights in the book, but not the copyright owner’s rights in the work. Using the term ‘possessor (or user) of the copy’ might therefore imply that the user innovation theory of the numerus clausus justifies ‘efficient theft’ (which it does not). Efficient theft is a highly controversial concept in private law, see e.g. Daniel Friedmann, The Efficient Breach Fallacy, 18 J. LEGAL STUDIES 1, 4 (1989). It is also true that an individual can violate the rights in a work without violating the rights in a copy – e.g., if I look over your shoulder and copy your original work of visual art, I have not violated your rights in the copy, but I have violated your copyright in your original work. Alternatively, I can violate neither the rights in a work, nor the rights in a copy (whether I own, possess or merely use the copy), if fair use excuses my activity. For example, if I look over your shoulder, and create a visual parody of your original work of art, I have not violated your rights in the copy or your rights in your work. In other words, fair use regulates the use of the work, not the use of the copy, and so it applies to owners, possessors and users of copies.}\]

\[\text{\footnotesize[34 The owner of the copy does not get all rights in the copy, of course. Although there is no general use right in copyright law, several of the rights granted under 17 U.S.C. §§ 106(1)-(6) have the effect of limiting a copy owner’s ability to ‘use’ the copy. For example, the public performance right, supra note 2, limits the ability of}\

186
the *copy* reflects the fact that the rights to use and resell, enshrined in the venerable common law policy against restraints on alienation, are considered to be fundamental incidents of legal ownership.\(^{35}\) We discussed the right to use in Chapter 4; in this chapter we will focus our comments on the right to resell. In the case of chattels and land, the right to resell is rarely called into serious question by either courts or scholars.\(^{36}\) In the intellectual property arena, however, the right to resell ‘copies’ or ‘embodiments’ of intangible works is frequently challenged in the courts by intellectual property owners who wish to restrain resales, and in the literature by scholars who are unconvinced by the theoretical underpinnings of the policy. The intellectual property equivalent of the policy against restraints on alienation is referred to as the ‘first sale’ or ‘exhaustion’ doctrine.

\[i) \text{ First Sale}\]

The U.S. Copyright Act addresses the division of rights in works and copies via the ‘first sale’ doctrine.\(^{37}\) Codified in Section 109 of the Act, this provision establishes that even though a copyright owner has the right to distribute copies of the work of authorship, once title to any particular copy has been transferred, the copyright owner no longer has the right to restrict further distribution of that particular copy.\(^{38}\) This principle also appears in patent law, where it is more commonly referred to as the principle of ‘exhaustion’, because the intellectual property owner’s right to control distribution of the patented item are literally ‘exhausted’ once title to that item has been lawfully transferred.\(^{39}\) In some ways the term exhaustion is more accurate than the term first sale, because transfer of title to a copy can be accomplished in ways other than by sale; nevertheless, the terms are generally used interchangeably.

\[^{35}\text{See e.g. Tony Honoré, Ownership, in Making Laws Bind 165 (Clarendon Press 1987): “Ownership comprises the right to possess, the right to use, the right to manage, the right to the income of the thing, the right to the capital, the right to security, the rights or incidents of transmissibility and absence of term….”}\]

\[^{36}\text{Though this is not always true. See e.g. Percy Bordwell, Alienability and Perpetuities II, 23 Iowa L Rev 1, 14 (1937); and Glen O. Robinson, Personal Property Servitudes, 71 U. Chi. L. Rev. 1449, 1480-81 (2004).}\]

\[^{37}\text{17 U.S.C. § 109(a).}\]

\[^{38}\text{Id.}\]

\[^{39}\text{See e.g. Quanta Computer, Inc. v. LG Electronics, Inc., 128 S. Ct. 2109, 2115 (2008): “The longstanding doctrine of patent exhaustion provides that the initial authorized sale of a patented item terminates all patent rights to that item.”}\]
Although copyright’s version of the first sale principle is codified in the Copyright Act, the doctrine’s origins are judicial rather than congressional. Its explicit entry into copyright law is generally traced back to the *Bobbs-Merrill* case, in which a publisher had sought to enforce a price maintenance scheme by means of a notice printed inside its books. The Supreme Court determined that a publisher could not prevent the resale of its books at less than a minimum price: “To add to the right of exclusive sale the authority to control all future retail sales, by a notice that such sales must be made at a fixed sum, would give a right not included in the terms of the statute, and, in our view, extend its operation, by construction, beyond its meaning, when interpreted with a view to ascertaining the legislative intent in its enactment.”

Having reached the conclusion that the statutory right to ‘vend’ copies only included the right to control the first sale of those copies, the court found it unnecessary to address the question of whether the notice printed inside of the books constituted an agreement, operated as a license, or provided sufficient notice to give rise to an equitable servitude. This omission was procedurally appropriate, as the case was argued on the basis of statutory interpretation alone; however, the court’s confidence in the intent of Congress was somewhat illusory. Even though Justice Day repeatedly stated that the right to vend could not be interpreted to include the ability to attach post-sale restrictions to copies of copyrighted works, this conclusion was reached more by way of mere assertion, than via a process of reasoning. Although he indicated that this was a matter of Congressional intent, there was

---

40 The notice read: "The price of this book at retail is one dollar net. No dealer is licensed to sell it at a less price, and a sale at a less price will be treated as an infringement of the copyright. THE BOBBS-MERRILL COMPANY." *Bobbs-Merrill Co. v. Straus*, 210 U.S. 339, 341 (1908). The publisher sold the book at a discount to wholesalers, and wanted to ensure that this discount did not put downward pressure on retail prices.

41 *Bobbs-Merrill*, *id.* at 351.

42 “[T]his is purely a question of statutory construction. There is no claim in this case of contract limitation, nor license agreement controlling the subsequent sales of the book. … This conclusion renders it unnecessary to discuss other questions noticed in the opinion in the Circuit Court of Appeals.” *Bobbs-Merrill*, supra note 40 at 350-51. It was proven on the record that the booksellers knew of the notice when it purchased the books. *Id.* at 342.

43 *Bobbs-Merrill*, supra note 40 at 350-351: “In our view the copyright statutes, while protecting the owner of the copyright in his right to multiply and sell his production, do not create the right to impose, by notice, such as is disclosed in this case, a limitation at which the book shall be sold at retail by future purchasers, with whom there is no privity of contract. This conclusion is reached in view of the language of the statute, read in the light of its main purpose to secure the right of multiplying copies of the work, a right which is the special creation of the statute. True, the statute also secures, to make this right of multiplication effectual, the sole right to vend copies of the book, the production of the author's thought and conception. The owner of the copyright in this case did sell copies of the book in quantities and at a price satisfactory to it. It has exercised the right to vend. What the complainant contends for embraces not only the right to sell the copies, but to qualify the title of a
no discussion of what Congress actually had intended, nor any evidence that it had even considered the question.\textsuperscript{44}

Thus, it must be admitted that the venerable pedigree of first sale rests upon a shadowy rationale which was so powerful that it could remain both unspoken and unexamined. It is generally accepted\textsuperscript{45} that this animating force was none other than the common law’s discomfort with restraints on alienation; thus, Bobbs-Merrill represented the importation of common law property principles into the ‘wholly statutory’\textsuperscript{46} arena of copyright.

Patent law has its own version of the first sale doctrine, which is usually referred to as ‘patent exhaustion.’\textsuperscript{47} Like copyright owners, patentees seek to avoid the application of the doctrine of exhaustion by classifying their transactions as licenses rather than sales. Despite these efforts, courts continue to restrain the ability of sellers to control the use and retransfer of patented items after they have been physically transferred to purchasers. Consider the most recent Supreme Court pronouncement on the issue: “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.”\textsuperscript{48} This ruling in \textit{Quanta v. LG} prevented the patent owner from 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\textsuperscript{49}), we can see why the patent owner would want to capture that downstream value.

\begin{quote}
\textsuperscript{44} This has caused some scholars to criticize the case, e.g., Robinson, \textit{supra} note 36 at 1470-71.
\textsuperscript{45} See e.g. Melville B. Nimmer and David Nimmer, 2 Nimmer on Copyright § 8.12(A) (Matthew Bender 2002).
\textsuperscript{46} \textit{Bobbs-Merrill, supra} note 40 at 346. Congress later codified the result in Bobbs-Merrill in the 1909 Act, and carried it forward in the 1976 Act.
\textsuperscript{47} \textit{Quanta, supra} note 39.
\textsuperscript{48} \textit{Id.} at 2122.
\textsuperscript{49} 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.
\end{quote}
In terms of its pronouncements on the principle of exhaustion, the Quanta case represents a straightforward affirmation of 150 years of patent jurisprudence. Its doctrinal significance goes far beyond this however, because the case dealt with the application of the doctrine of exhaustion to method patents. Method patents literally claim a method of achieving a result - a process. Thus, unlike with a patented device, there is no embodiment within the claims of a method patent itself. 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 embodiment of a patent. Nonetheless, the Supreme Court decided that once a method claim is substantially embodied in any particular device and that device is sold, then the exhaustion doctrine takes effect. Once this happens, the patent owner cannot demand payments or require permission for further use or transfer.

ii) Criticisms of First Sale/Restraints on Alienation

Both legal and economic commentators agree that the right to sell is one of the fundamental aspects of property ownership. However, once a property right is initially allocated to a particular owner, the right to divide rights via contract (which is also considered to be a fundamental aspect of property ownership), raises the possibility that the right to contract will itself be used to restrict the right to resell. In other words, once property rights are allocated, we are then faced with the conflict between the property and contract rights of the initial owner, and the property rights of the downstream transferee. In the case

50 “For over 150 years this Court has applied the doctrine of patent exhaustion to limit the patent rights that survive the initial authorized sale of a patented item.” Quanta, supra note 39 at 2113.

51 Quanta, supra note 39 at 2117: “Nothing in this Court's approach to patent exhaustion supports LGE's argument that method patents cannot be exhausted. It is true that a patented method may not be sold in the same way as an article or device, but methods nonetheless may be "embodied" in a product, the sale of which exhausts patent rights. Our precedents do not differentiate transactions involving embodiments of patented methods or processes from those involving patented apparatuses or materials. To the contrary, this Court has repeatedly held that method patents were exhausted by the sale of an item that embodied the method. In Ethyl Gasoline Corp. v. United States, for example, the Court held that the sale of a motor fuel produced under one patent also exhausted the patent for a method of using the fuel in combustion motors. Similarly, as previously described, Univis held that the sale of optical lens blanks that partially practiced a patent exhausted the method patents that were not completely practiced until the blanks were ground into lenses.” (Citations removed.)

52 Id.

53 Id.

54 See e.g. Hohfeld, supra note 1 (legal powers); and Yoram Barzel, Economic Analysis of Property Rights 114 (2d ed. 1997) (ability to transfer).

55 See e.g. Steven N.S. Cheung, The Structure of a Contract and The Theory of a Non-Exclusive Resource, 13 J. L. Econ. 49 (1970).
of intellectual property copies, this conflict is discussed in terms of the first sale doctrine; while in the case of chattels and land it is a subset of the servitude issue.\textsuperscript{56}

There are several criticisms of the first sale doctrine, which can be divided into roughly three categories: 1) criticisms that relate to the conflicts generated when different modes of law and equity seek to regulate rights (i.e., freedom of contract, the priority of property rights, and the impact of notice on both contract and property law); 2) criticisms based on arguments about economic efficiency; and 3) procedural criticisms of the case law that imported the first sale doctrine into intellectual property law. Because the criticisms in this first category apply to both the first sale doctrine and to the policy against novel servitudes, we discussed arguments from this first category when we examined servitudes in Chapter 4.\textsuperscript{57} In this section, we will only deal with criticisms that relate to economic efficiency and to the case law foundations of the first sale doctrine.\textsuperscript{58}

Arguments about the economic efficiency of first sale center around the doctrine’s interference with differential pricing or price discrimination: “Since the effect of the doctrine is that the copyright owner has only one opportunity to earn revenues from the sale of a particular copy, the price of that copy on its initial sale will usually be higher – and thus out of reach for some prospective buyers – than it would be if the copyright owner were also able to earn revenues from each successive sale or loan of the copy.”\textsuperscript{59} In other words, if purchasers can be prevented from reselling and lending their own copies, then everyone will have to buy or rent a copy directly from the copyright owner. Copyright owners can then sell copies at a higher price to those who value the good at above the single market clearing price; and at a lower price to those who value the good at more than its marginal cost of production, but less than its single market clearing price. Relative to uniform pricing under conditions of monopoly, price discrimination is a boon for producers, consumers and society. If a copyright owner can price discriminate, higher-valuing purchasers subsidize lower-valuing purchasers, and society is better off because everyone who values the good at above its marginal cost can now purchase the good. Price discrimination avoids the deadweight loss

\textsuperscript{56} We discussed the servitude issue in Chapter 4, Part 1, \textit{The Legal Policy Against the Creation of Novel Servitudes}.

\textsuperscript{57} \textit{Id.}

\textsuperscript{58} This is simply to avoid duplication.

\textsuperscript{59} \textsc{Paul Goldstein}, \textsc{Copyright’s Highway: From Gutenberg to the Celestial Jukebox} 201 (Rev. ed. 1994).
of monopoly pricing because output under differential pricing is greater than that under monopoly pricing.\(^\text{60}\)

Price discrimination sounds like a win-win situation, until we consider a critical economic caveat: price discrimination alleviates the deadweight loss of monopoly pricing. Thus, if there is no monopoly – i.e., if there are other suppliers of the good - then competition will increase supply and decrease prices to socially optimal levels.\(^\text{61}\) This is exactly what happens (albeit in a limited way) when a purchaser of a good becomes a reseller of that same good: “The first sale doctrine fosters indirect competition by allowing the first buyer of a work to resell it. Resale is especially detrimental to price discrimination if a low valuation user purchases a work at a low price and arbitrages by selling the work to a high valuation user.”\(^\text{62}\) Thus, the existence of the first sale doctrine itself compromises the monopoly of the copyright owner, at least in regards to the supply of lawfully made copies. And this means that, with the first sale doctrine in place, there may literally be no problem for price discrimination to solve.\(^\text{63}\)

A second set of criticisms of the first sale doctrine are aimed at the case that created the doctrine (the Bobbs-Merrill case), rather than at the concept itself. We have already seen that that court’s confident reliance on the intention of Congress was rather illusory.\(^\text{64}\) But an even more unusual fact about the Bobbs-Merrill case is that although it is celebrated for importing the common law policy against restraints on alienation into copyright law, the case itself did not involve a restraint on alienation. The publisher in Bobbs-Merrill did not seek to prevent the actual resale of its books; it merely wanted their retail prices to remain above a certain level.\(^\text{65}\) Many first sale disputes are of this price maintenance variety, and in these and many other cases, the existence of ‘restraints on alienation’ is more assumed than actual.\(^\text{66}\) That is, the judiciary often assumes that the restriction that the original seller seeks to impose will indirectly restrain further alienation of the item (by making the item less attractive to

---

\(^\text{60}\) Deadweight loss occurs when buyers, who value the good at above marginal cost, but below the uniform monopoly price, do not buy the good.


\(^\text{63}\) Gordon, supra note 61 at 1386-87.

\(^\text{64}\) Bobbs-Merrill, supra note 40.

\(^\text{65}\) Robinson, supra note 36 at 1470.

\(^\text{66}\) Id.
prospective buyers), even though the restriction does not prevent resale itself. This discontinuity between doctrine and empirical fact has been noticed by scholars, and has provided the fodder for much criticism.\(^\text{67}\)

Despite the accuracy of this criticism with respect to the Bobbs-Merrill case itself, there are situations in which sellers do wish to prevent actual resale of an asset. In copyright law, many of them involve the redistribution of promotional copies.\(^\text{68}\) Other copyright owners are motivated by a desire to prevent resale because resale is seen as a guise for copyright infringement. If a copy can be cheaply and easily copied, then an individual might resell a lawfully purchased copy to someone else, but keep a copy of it for himself. This might happen with CD’s for example, but is less likely to happen with bound books. Other copyright owners simply want to restrict the formation of a market in used copies, because the ability to sell an original copy to each and every purchaser improves the copyright owner’s bottom line. Restrictions on actual resale also occur in the case of patented goods. For example, patent owners might try to restrict the resale of a patented item because they do not want their patented components combined with the components of other manufacturers.\(^\text{69}\)

Clearly, the ability to restrict resale is advantageous for sellers, because all buyers must purchase goods from the original seller. Of course, what is beneficial for the original seller of an asset is disadvantageous for resellers interested in developing a market in used goods. But does a restriction on resale also have a deleterious impact on user innovation? In other words, can the legal solicitude for the right to resale be justified in terms of user innovation? We examine this question in detail in the next section.

iii) The ‘Mere’ Right to Rent/Resell Fosters User Innovation

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 prohibit use restrictions on assets. But the ‘mere’ right to sell or temporarily transfer an asset without the permission of the original owner (or producer) doesn’t seem to have anything to do with innovation per se. The case of

\(^{67}\) Id.
\(^{68}\) A recent example of one such case was discussed in Chapter 4, Part 2, *IP Copies and Notice: The Realities of the Transaction*.
\(^{69}\) This happened in the Quanta case, *supra* note 39. Intel did not want its components combined with non-Intel components.
patented components is perhaps an exception to this, because restrictions on resale are often used to prevent the recombination of components in new ways,\textsuperscript{70} and this can obviously have an adverse impact on the development of improved products.

But what about those assets which are not generally used as components of finished products, but are sold as-is to the end consumer? In these cases, the ability to resell would appear to be entirely non-innovative, and would appear to impact only the distribution of wealth (the way the pie is divided), and not the amount 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 we will go beyond this observation and take a look at how the law’s general refusal to enforce resale restrictions 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 machines and from Internet companies who deliver them by mail. 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.\textsuperscript{71}

There were many reasons for this. Many Hollywood executives were simply of the opinion that “rental wouldn’t work.”\textsuperscript{72} 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.\textsuperscript{73} They were willing to sell high-priced VHS cassettes to the serious movie buff, but they didn’t want these

\textsuperscript{70} Id.
\textsuperscript{71} For a detailed history of the VCR and related technologies, see generally FREDERICK WASSER, VENI, VIDI, VIDEO 113 (University of Texas Press 2001) [Wasser].
\textsuperscript{72} Id. at 113.
\textsuperscript{73} Id. at v.
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 thwarted 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: this was to prevent distributors from renting, rather than selling, the tapes to end users. Of course, the weakness in this scheme was that the purchaser of the tapes could simply ignore the contractual restrictions on the distributor and set up shop as video rental outlets. Other studios therefore sought to develop exclusive partnerships and ‘lease-only’ deals with the rental outlets themselves. If tapes were leased rather than sold to retailers, then studios could claim a portion of every rental transaction. In short, the “majors wanted to either sell directly to the customer or participate in each and every rental transaction – in other words, receive a cut of the rental fee.”

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 U.S. 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, 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

74 Id. at 101.
75 Id. at 62.
76 Id. at 64.
77 Id. at 111.
78 Id. at 151.
79 Id. at 101.
80 There were other factors that contributed to Cartrivision’s failure. Cassettes were kept in stock at the point of rental but instead had to be pre-ordered by the customer wishing to rent the tape. This made the process of rental cumbersome. In addition, the entire stock of cassettes decayed in the warehouse just before Christmas of 1972, which likely provided the death blow to an already shaky strategy. See Michael Raynor, The Strategy Paradox 21 (Broadway Business 2007).
81 The first sale doctrine is discussed supra notes 37-46 and accompanying text.
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 to end users. 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. With the failure of all these strategies, the only thing left for Hollywood to do was to charge more for the original sale of the cassette. 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, at a time 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.

In case there are any 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” wasser at 115) 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 a significant number of video rental outlets in the U.S., so members of Congress would have been faced with many angry constituents if they had given Hollywood control over movie rental so late 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.

Wasser, supra note 71 at 115.

Id. at 116.

Id. at 104-130.
This example is also interesting because it shows us how the law, by allowing users to ignore the governance rule that Hollywood sought to impose on VHS cassettes, facilitated the development of that innovative business model. And video rental had other downstream effects which no one could have predicted. Although the early days of video rental had provided independent film producers with a critical source of revenue, allowing them to enter new markets and expand production, in the end home video would help fuel the move toward industry consolidation and the movement towards the blockbuster model of movie production.

Early on, when the major studios were slow to release titles on video and were antagonizing the rental outlets by attempting to interfere with their business model and muscle in on their profits, the availability of independent titles was a boon to rental outlets and consumers alike. And in those early days, the consumer’s appetite for video rental was so strong that people would rent any available title, even obscure and low budget options. Thus, video rental became a major source of revenue for independents, through the sale of cassettes to retailers, consumers and rental entrepreneurs. As the market matured, however, sales driven by video rental outlets became a major source of revenue for the major studios, gradually surpassing box office rental in fiscal importance. In 1980 for example, studio income from the box office represented 98.05 percent of negative costs, while home video returns were only 0.79 per cent. By 1993, those figures had changed dramatically, with box office rental income declining to 46.37 per cent of negative costs, and home video returns rising to 93.47 per cent. However, this was not a case of cannibalization, because “[t]he value of the theatrical market had remained constant in absolute terms and had declined in relative terms.” What this means is that returns from home video provided a new source of revenue, which supplemented, rather than replaced, theatrical revenue. These swelling coffers financed the movement towards the blockbuster model of movie production, in which enormous production and advertising budgets funded the production of a small number of

---

86 A film’s ‘negative cost’ is the cost of producing a completed film negative. This includes the actual costs of production, but is somewhat broader. A film’s negative cost includes all costs related to rights acquisition, preproduction, production, postproduction and insurance. It may also include other costs, but it does not include distribution and exhibition costs. A film ‘breaks even’ when its revenues equal its negative cost. For a helpful discussion of negative cost (and of the difficulty in calculating it), see JOHN W. CONES, THE FEATURE FILM DISTRIBUTION DEAL 82-85 (Southern Illinois University Press 1997).
87 Wasser, supra note 71 at 171.
88 Id. at 170.
titles, which could be pushed through to the video market in high volumes thanks to theatrical and advertising exposure. Notice the synergy here: home video returns financed big budget movies and multi-outlet advertising, which in turn provided the consumer awareness and shelf space to drive video sales and rental. Under these conditions, independents found it increasingly difficult to compete, both in the production of content and in the competition for consumer attention. Over time, home video returns thus helped to weed out the independents that it had helped create, resulting in a high degree of consolidation in the movie industry.

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 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 useful feature of 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.)

iv) The Future of First Sale

In the past, the first sale doctrine worked well to apportion rights between copyright owners and lawful purchasers of copies of their works. When it was relatively cumbersome and expensive to make copies of intellectual property works, the division of rights between

---

89 Id. at 169.
the owner of the copyrighted work (e.g., a literary work), and the owner of a copy of the copyrighted work (e.g., a book), effectively created a regime in which the copyrighted work was akin to a parcel of land, and the copy of the work was akin to a chattel. This legal treatment makes sense from the point of view of user innovation because copies of works are functionally much like chattels: they are non-unique, movable material objects which are generally produced and distributed in large numbers. As with chattels, a great deal of innovation emerges from the use and modification of copies; therefore, giving the owners of copies the ability to use and to manipulate them in certain ways makes sense (though these abilities are not as absolute as they are in the case of traditional chattels).

There have always been difficulties with the differential legal treatment of works and copies, of course. By definition, every ‘copy’ embodies a copyrighted work, so there is no way to completely separate the use of the ‘work’ from the use of the ‘copy’. For example, when I frame a copyrighted poster, am I simply using my copy, or am I creating a new derivative work? What if I take a card and frame or mount it so that it looks like a poster? And even when it was technologically difficult for users to copy and modify works, it was still possible for users to perform them. When I perform a work by reading it out loud, am I making a use of the ‘work’ or the ‘copy’? It would seem both. Copyright law largely solved this problem by distinguishing between ‘public’ and ‘private’ uses of the work and its copies. A private performance was reserved to the owner of a copy, because a private use is more like a use of the chattel (the copy). After all, why would I purchase an album if I could not make its sounds audible in my living room? A public performance, on the other hand, was reserved to the owner of the work, because a public use seems more like a use of the ‘work’. Imagine playing that album in a nightclub, for example, or performing a play on a Broadway stage – in these cases the value of the use seems to derive more from the work itself rather than the mere copy.

But it was the development of copying technology that began to seriously erode the differential legal treatment of works and copies, because users could easily make new copies of works, in addition to using and modifying the copy that they had lawfully purchased.

---

90 This is true only during the term of copyright protection, of course. Once a copyright expires, then even the work itself is available for everyone to freely copy and modify.
Digitization of works, and the ability to trade files over networked computers, eroded it further. However, not all of the mischief of technology was at the expense of the owners of copyrighted works. Indeed, with technological developments came the ability to use Digital Rights Management (DRM) to control a user’s ability to use and interact with copies of works in ways that the Copyright Act had always taken for granted. For example, an owner of a copyrighted work might want to limit the number of times a user can view a copy of a work, by encoding it with an automatic expiry system. Or the owner of a copyrighted work might want to restrict even that small amount of user modification and copying that was traditionally allowed via the fair use defense. Thus, DRM gives owners of works the ability to control the work via direct control of copies themselves,\(^{92}\) and so this is where the traditional distinction between an intellectual property work and a copy of that work truly breaks down.

The problems that technology causes for the categorization of ‘works’ and ‘copies’ are not unique to intellectual property assets, however. Similar problems also occur with tangible assets. When sellers seek to restrict the use and modification of a tangible asset, technology interferes with the law’s ability to classify a transaction as a ‘contract’ or ‘notice’ or ‘property’ transaction. In Chapter 4, we saw how this act of legal classification enabled the law to regulate the use of the three legal mechanisms of governance – contract, notice and property. Regulating the use and availability of these mechanisms of governance was key to preserving user innovation. Now, we examine how the law should respond to what is for sellers the potentially most powerful and cost-effective mechanism of governance: technology.

**PART 2 \ THE FUTURE OF THE NUMERUS CLAUSUS - THE CHALLENGE OF TECHNOLOGY**

Not surprisingly, the law has been far more adept at dealing with the legal mechanisms of governance - contract, notice and property - than it has been at dealing with

---

\(^{92}\) This control is not perfect, of course. Even if a user is not technically savvy enough to circumvent DRM, she can sometimes make a manual copy. For example, even if I cannot directly copy a DRM restricted e-book, I can view it and retype the words on my computer. With textual material, this might be a close enough substitute for users, but with audiovisual works, such low-tech ways of copying are poor substitutes.
technology. The technological mechanism of governance has proven to be particularly troublesome for the law, because its extra-legal nature can act as either a supplement to, or a replacement for, the law.\textsuperscript{93} Technology supplements the law when it impacts the legal categorization of a transaction. For example, the need to actively click on an ‘I Agree’ icon when installing computer software might be a manifestation of assent, and so give rise to contractual obligations. On the other hand, if the computer software is programmed to shut down if the consumer uses the product in violation of asserted terms, then the ability of the law to regulate this act of governance is circumvented altogether.\textsuperscript{94}

Technological pre-emption of the law provides just one of the reasons that technology is proving to be so challenging to the law’s regulation of governance. Another is that technology interferes with the methods that the law has traditionally used to draw the line between governance and exclusion. Technology interferes with the law’s ability to regulate governance via asset type, because technology can be used to change the asset itself. When a manufacturer encodes use restrictions into a software program, is this implementation of a governance rule, or just product design? Whether we categorize it as the former or the latter has important implications for the degree of legal interference. Technology also interferes with the law’s ability to regulate governance via transaction characteristics, because it blurs the line between notice and contract, between labeling and bargaining. When I have to click through a series of license terms when installing software, is this mere notice, or does the act of clicking manifest sufficient contractual assent? Indeed, it can be said that technology blurs the line between the asset itself and the transaction involving the asset. While this is not the first time that the line between contract and property has been blurred,\textsuperscript{95} technology is proving to be far more challenging to legal theory than standard form contracts ever were.

The judiciary has always been keenly aware of the challenge that notice poses for the law’s attempt to place limits on transactions involving property, and they also knew that this

\textsuperscript{93} To complicate matters further, the law is sometimes used to enforce the technological mechanism of governance. \textit{See infra}, \textit{Digital Rights Management and the Digital Millenium Copyright Act}.

\textsuperscript{94} \textit{See e.g.} Margaret J. Radin, \textit{Regulation by Contract, Regulation by Machine}, 160 J. INSTITUTIONAL & THEORETICAL ECONOMICS 142 (2004) at 160: “The advent of Digital Rights Management Systems (DRMS’s) has the potential to read out the regulatory contouring of contract just as the advent of [technologically implemented contracts]… has the potential to read out the regulatory contouring of [intellectual] property.”

\textsuperscript{95} \textit{See e.g.} Arthur Allen Leff, \textit{Contract as Thing}, 19 AM. U. L. REV. 132 (1970). The technological ability to move seamlessly and costlessly from exclusion to governance raises the possibility of what might be called ‘Thing as Contract’ – i.e., a world in which things are not just replaced by contracts, but where things \textit{act} like contracts.
challenge would only increase as the means to provide notice grew easier and cheaper: “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 burthens 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 technology 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?

But technology is not just troublesome for legal doctrine and theory. Because technology can actively prevent users from using and modifying copies in ways that were traditionally allowed by the Copyright Act, it also poses a grave potential threat to user innovation. As we saw in Chapter 2, the right to modify an asset has the most direct impact on user innovation, because users often modify existing assets to create new and improved products. Manufacturers use two principal mechanisms to prevent users from modifying the assets that they possess.\(^\text{96}\) The first mechanism, Digital Rights Management (DRM) or Technological Protection Measures (TPM), is technological.\(^\text{97}\) The second, the Digital Millennium Copyright Act (DMCA), adds an extra layer of legal protection to the technological measures.\(^\text{98}\) Because DRM and the DMCA are primarily used in conjunction with intellectual property products, and because they work synergistically to interfere with the process of user innovation, we will discuss them together.

\(^\text{96}\) Of course, manufacturers also use other methods to interfere with the ability to modify assets. Product warranties are often made conditional on the non-modification of an asset, particularly with durable products that have mechanical or electronic components that might break down or malfunction early in the product life cycle. This method of interfering with the right to modify an asset is not considered to be a substantial threat to user innovation, because users can simply choose between warranty protection and asset modification. And since users’ modifications or attempts to repair a product may negatively impact its performance and make it more costly to return it to its original condition, it would be unreasonable for a manufacturer to have to provide warranty protection for such products.

\(^\text{97}\) There are many definitions for DRM and TPM. See e.g. Renato Iannella, Digital Rights Management (DRM) Architectures, D-Lib Magazine, Vol. 7 No. 6, June 2001, available at http://www.dlib.org/dlib/june01/iannella/06iannella.html: “DRM covers the description, identification, trading, protecting, monitoring and tracking of all forms of usages over both tangible and intangible assets including management of rights holders relationships. Additionally, it is important to note that DRM is the "digital management of rights" and not the "management of digital rights". That is, DRM manages all rights, not only the rights applicable to permissions over digital content.”

Currently, implementing a use restriction on an asset is most realistic in regards to intellectual property assets that are used in conjunction with networked devices, where the digital nature of copies, the ubiquity of network connectivity, and the availability of automated electronic financial transactions combine to offer the possibility of a seamless, invisible and frictionless system of monitoring of (and payment for) incremental product usage. With digital assets, such as software programs installed on our computers, use restrictions can even be monitored and enforced via remote product deactivation.\(^99\) In most cases, enforcing a use restriction on a tangible chattel is unwieldy, because it requires a seller to keep track of a buyer’s use of the asset. However, this does not mean that technology only poses a challenge to the legal regulation of intellectual property assets. To the extent that new technologies enable somewhat greater monitoring, enforcement and payment for incremental use of tangible assets, the specter of technological control has relevance even for garden variety chattels.

For example, Radio Frequency Identification tags (RFID tags), smaller than the head of a pin, can be placed within tangible assets and be used to track their location.\(^100\) Although RFID tags are currently most widely used to track inventory in stores, they can presently be

\(^99\) See e.g. Gary Edwards, \textit{Self-Help Repossession of Software: Should Repossession Be Available In Article 2b Of The UCC?}, 58 U. PIT. L. REV. 763, 778-79 (1997): “The most widely publicized case of electronic disabling is Revlon v. Logisticon, Inc.. Revlon had entered into an agreement with Logisticon for the design, development, and installation of a software package to aid in inventory control. The project was to proceed in phases and a milestone payment was to be made at completion of each phase. To aid in development, Logisticon was allowed dial-up access to the Revlon computer system. Revlon was dissatisfied with the software and withheld $180,000 from one of the payments. Logisticon was apparently unsuccessful at repairing the program to Revlon’s satisfaction. After several months, Logisticon sent written notice of its intention to repossess the software. On the night the notice was sent, Logisticon dialed into Revlon’s system and electronically disabled the software.” (Citations removed.)

\(^100\) See e.g. BILL GLOVER AND HIMANSHU BHATT, RFID ESSENTIALS (O’Reilly Media 2006) at 1: “RFID stands for Radio Frequency Identification, a term that describes any system of identification wherein an electronic device that uses radio frequency or magnetic field variations is attached to an item. The two most talked-about components of an RFID system are the tag, which is the identification device attached to the item we want to track, and the reader, which is a device that can recognize the presence of RFID tags and read the information stored on them.” In 2007, one group of industry analysts stated that: “In 10 to 15 years RFID technology will be ubiquitous.” See V. DANIEL HUNT, ALBERT PUGLIA AND MIKE PUGLIA, RFID – A GUIDE TO RADIO FREQUENCY IDENTIFICATION (Wiley 2007) at 111. Currently, RFID’s principle areas of application are in: tracking systems (animals, individuals, vehicles, inventory); payment systems (particularly point of sale purchase); access systems (such as key cards); and control of counterfeit goods. \textit{Id.} at 1 and 73-74. The smallest RFID chip currently in commercial production, which can be embedded in a piece of paper, is 0.4x0.4 mm. See Hitachi Europe Limited, \textit{Mu Chip Data Sheet}, available at \url{http://www.hitachi-eu.com/mu/Products/Mu%20Chip%20Data%20Sheet.pdf}. (Last visited June 9, 2010.) Hitachi has also developed a smaller chip, 0.05x0.05mm, which is so small that it looks like ‘powder’ to the human eye: RFID ‘Powder’ – World’s Smallest RFID Tag, available at \url{http://www.technovelgy.com/ct/Science-Fiction-News.asp?NewsNum=939}. (Last visited June 9, 2010.)
used to track and levy charges for continued use of tangible assets, albeit in a blunt way. Library books, for example, are already tracked and charged for with the help of RFID tags. As such tracking technologies become more sophisticated and cheaper to implement, what is a current reality for intellectual property copies may become somewhat more realistic for tangible chattels as well. Indeed, even technological developments that are not specifically related to asset tracking can sometimes facilitate the enforcement of servitudes and other asset restrictions. Recall the UMG case\textsuperscript{101} and the dispute between Laura Flores and Livingsoft,\textsuperscript{102} for example. In both of these cases, the Internet and the eBay service worked together to make it easier for sellers to detect when purchasers were violating restrictions on resale, even though neither of these technologies was specifically designed for this purpose.

In addition to developments in law (e.g., the DMCA), and developments in technology (everything from tracking technologies to general purpose technologies like the Internet), the need to address the relationship between the \textit{numerus clausus} and technology also arises from developments in legal and economic theory. Previous theories of the \textit{numerus clausus} point to an inverse relationship between the \textit{numerus clausus} and the ability of technological developments to facilitate notice. The following quote from Merrill and Smith exemplifies this view: “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. Just as the rise of land registers allowed some loosening of the \textit{numerus clausus}, so too technology that lowers information costs can be expected to weaken the \textit{numerus clausus} further.”\textsuperscript{103} As we shall see, the user innovation theory is far less sanguine about the wisdom of weakening the \textit{numerus clausus} in the face of technological developments.

In section 1, we examine DRM and the specialized legal regime created by the DMCA. In section 2, we ask whether the presence of technological controls interferes with the development of feelings of psychological ownership, because this might warrant a loosening of the \textit{numerus clausus}. Finally, in section 3, we discuss the appropriate legal response to technological controls, and we also discuss possible business alternatives to servitude regimes and technological restrictions.

\textsuperscript{101} Infra Chapter 4, Part 2, \textit{IP Copies and Notice: The Realities of the Transaction}.
\textsuperscript{102} Infra Chapter 3, Part 2, \textit{The Impact of Market and Social Norms on the Willingness to Freely Give Permission}.
1) Digital Rights Management and the Digital Millennium Copyright Act

It is true that copyright law gave the owners of copyrighted works the right to interfere with user innovation long before the enactment of the DMCA. One of the most significant methods of interfering with user innovation was the granting of a derivative work right, which gave the owner of a copyrighted work the right to prepare additional works based upon that initial work. However, the derivative work right has traditionally been limited by the fair use defense (which gives users the ability to creative derivative works of certain types - parodies, for example), and by the sheer difficulty and expense of bringing a lawsuit. With DRM and the DMCA, in contrast, the owner of the copyrighted work has the ability to use technology and the law to prevent the modification of copies and works even before courts have a chance to intervene.

DRM can be defined as any technological measure which provides a way for intellectual property owners to directly protect and enforce their rights. An example of DRM would be a copy control on a DVD or CD which prevents the disc from being copied.

---

104 A derivative work is defined in 17 U.S.C. § 101: ‘A “derivative work” is a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted. A work consisting of editorial revisions, annotations, elaborations, or other modifications which, as a whole, represent an original work of authorship, is a “derivative work”.’

105 The derivative work right did not always exist. It was added after some courts determined that the right to reproduce a copyrighted work in copies did not include such things as translating a book into another language. See Stowe v. Thomas, 23 F. Cas. 201, 207 (C.C.E.D. Pa. 1853): “A “copy” of a book must, therefore be a transcript of the language in which the conceptions of the author are clothed; of something printed and embodied in a tangible shape. The same conceptions clothed in another language cannot constitute the same composition, nor can it be called a transcript or “copy” of the same “book.” The derivative work right was added to the Copyright Act in 1870.

106 Actually, DRM allows intellectual property owners to protect their property in ways that go beyond the rights set out in the Copyright and Patent Acts. See e.g. Julie Cohen, Copyright and the Jurisprudence of Self-Help, 13 BERKELEY TECH. L.J. 1089, 1094-95 (1998): “Unhappily for consumers, however, digital rights management regimes will enable information providers to appropriate far more protection against copying and distribution than intellectual property law now provides. Copyright law allows some reuse of protected expression under a variety of exceptions designed to serve the public interest, and allows any reuse after the term of copyright protection has expired. Copyright also does not attach to facts, ideas, or functional principles; instead, it treats these materials as public domain "building blocks" for future works. Many compilations of information consist largely of such public domain material and are protected only minimally, if at all, by copyright. The common law tort of data misappropriation provides some protection for uncopyrightable facts, but cannot protect against all copying, since the Copyright Act expressly preempts state-created rights that are "equivalent" to the rights afforded under copyright law. Trade secrecy law protects only information that is not generally known or readily ascertainable, and allows discovery of protected information by reverse engineering and other "proper means." In short, legal protection against unauthorized copying and distribution is incomplete, and is so by design.” (Citations removed.)
Both DRM and the DMCA are motivated by the belief that digital media and electronic distribution make large scale piracy so cheap and easy that such activity needs to be dealt with even *before* it occurs. After all, if I can easily copy the files on a DVD or CD and distribute them via the Internet, it may be virtually impossible for the copyright owner to stop the swelling tide of highly distributed copying that results.

DRM adds an *ex ante* layer of technological protection to the *ex post* layer of legal rights given to copyright owners by the Copyright Act. In turn, the DMCA adds yet *another* layer of legal rights on top of the technological measures. To accomplish this, the DMCA takes a bipartite approach: 1) it makes certain acts of circumventing DRM unlawful; and 2) it prohibits the manufacturing and distribution of tools that are used to circumvent DRM. Violations of the DMCA exist *in addition to* any underlying violations of copyright law which may or may not take place. The DMCA focuses on tools that circumvent, in addition to acts of circumvention, in order to provide another layer of *de facto* protection for copyright owners. After all, many users may be able to use a tool to circumvent a DRM measure, but may lack the technological know-how to actually create such a tool for themselves. Thus, targeting tool manufacturers stops some acts of circumvention even before they occur.

Together, DRM and the DMCA interfere with user innovation in several ways: 1) they make it legally risky to circumvent a technological protection measure; 107 2) they make it more difficult for users to get tools to circumvent technological controls; 108 3) they make activities that may not themselves violate copyright law unlawful; 109 and 4) they may chill activity that is *lawful* under both the DMCA and the Copyright Act because users and tools manufacturers may be worried about the uncertain parameters of liability. 110 Although technology provides intellectual property owners with new ways to control the downstream use of its products, technology also expands the ability of users to manipulate and modify these products in new and innovative ways. In many cases, copyright owners have responded

107 17 U.S.C. § 1201(a)(1)(A): “No person shall circumvent a technological measure that effectively controls access to a work protected under this title.”

108 17 U.S.C. § 1201(a)(2) and 1201(b)(1): “No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof, that” circumvents a technological protection measure.

109 See *e.g.* Sony Computer Entm't Am., Inc. v. GameMasters, 87 F. Supp. 2d 976 (N. D. Cal. 1999) (successful DMCA claim against region coding system where there was no violation of the copyright in game itself).

110 See *e.g.* Electronic Frontier Foundation, *Unintended Consequences: 12 years Under the DMCA*, February 2010, available at [http://www.eff.org/files/eff-unintended-consequences-12-years.pdf](http://www.eff.org/files/eff-unintended-consequences-12-years.pdf) at 2-4 [EFF].
to these new sources of innovation by filing suit under the DMCA. DMCA lawsuits have been filed over things as diverse as game skins, emulation software, dance moves for a dog robot, a free Internet game server, and computer chips that enabled laser cartridges to be refilled. In many of these cases, the DMCA lawsuits resulted in either a decision in favor of the copyright owner (the free Internet game server), the withdrawal of the innovation because the innovator could not afford the cost of litigation (the emulation software), or withdrawal of the innovation and the payment of a settlement fee (the free Internet game server). In other cases, the copyright owner either folded in the face of public outrage (the dance moves for the robot), or the courts refused to allow the DMCA to interfere with the innovation - but only after expensive and protracted litigation (the laser cartridge chips).

Although the DMCA has had a negative impact on innovation (both user and manufacturer), there are some ways in which the DMCA itself tries to mitigate this impact. First, in regards to acts of circumvention, the DMCA differentiates between the circumvention of access controls and the circumventions of copy controls. The act of circumventing an access control is unlawful, while the act of circumventing a copy control is not. If a user has a lawful copy of a DVD, for example, we want that user to be able to copy sufficient material to make fair uses of the copyrighted work. However, in practice this distinction has proven ineffective because it is easy to design DRM such that a single technological protection measure controls both access and copying. In most cases, then, every act of circumventing a copy control (which is not unlawful) simultaneously circumvents an access control (which is unlawful).

Second, there is a ‘savings clause’ which states that “Nothing in this section shall affect rights… or defenses to copyright infringement, including fair use.” This clause would seem to suggest that if an act of circumvention is undertaken, or if a device is used, for

---

111 Some of these innovations were created by users, others by manufacturers.
112 EFF supra note 110 at 14. A ‘skin’ improves the appearance of a character in a computer game, for example by giving it a new outfit, or removing its outfit entirely.
113 Id. at 15.
114 Id. at 16.
115 Id. at 16-17.
118 Id.
the purposes of making fair use of a copyright work, such an act or device will not violate the DMCA. However, to date, the courts have read this provision narrowly, such that it does not, in most cases, immunize circumvention acts or tools even where there is no underlying violation of copyright law.120

Third, there is a ‘rulemaking process’, via which the Register of Copyrights and the Librarian of Congress hold an extensive set of public consultations every three years, and can issue exemptions which allow users to circumvent access controls free of DMCA liability. However, the rulemaking process has turned out to be fundamentally antithetical to the process of user innovation. By its very nature, we saw that innovation is an emergent process: we cannot predict who will innovate, or how they will innovate. This means that the rulemaking process, which has so far demanded that there be both a well-defined, extremely specific class of users and class of uses, is simply not capable of recognizing (and protecting) the myriad activities and users that generate user innovation. In the 2003 rulemaking, for example, various proposed exemptions designed to allow users to make fair uses of copyrighted works recorded on DVDs were rejected,121 but a very narrow exemption for film and media studies professors for the purposes of classroom educational use was finally allowed in 2006.122 Although this exemption was expanded in the most recent rulemaking, and now applies to educational uses by college and university professors and film and media

120 The distribution of DeCSS, for example, which was designed to circumvent CSS on lawfully obtained DVDs, was not immunized by 1201(c) even though a predominant use of DeCSS was to make the DVDs playable on LINUX computers (which would normally be a fair use). See Universal City Studios Inc. v. Corley, 273 F.3d 429, 443 (2d Cir. 2001): “We disagree that subsection 1201(c)(1) permits such a reading. Instead, it simply clarifies that the DMCA targets the circumvention of digital walls guarding copyrighted material (and trafficking in circumvention tools), but does not concern itself with the use of those materials after circumvention has occurred. Subsection 1201(c)(1) ensures that the DMCA is not read to prohibit the "fair use" of information just because that information was obtained in a manner made illegal by the DMCA. The Appellants' much more expansive interpretation of subsection 1201(c)(1) is not only outside the range of plausible readings of the provision, but is also clearly refuted by the statute's legislative history.” In cases which involve products which allow fair uses of works traditionally covered by copyright - movies and software, for example - 1201(c) is unlikely to come to the aid of device manufacturers. However, in cases which involve products which do not seem to be related to copyright law (even though they might be embedded with computer chips that contain software code, such as printer cartridges and garage door openers) courts are more likely to read 1201(c) more robustly.
studies students, it still does not apply to other university students, nor to K-12 teachers.\footnote{Recommendation of the Register of Copyrights in RM 2008-8 June 11, 2010, pages 63-64 \url{http://www.copyright.gov/1201/2010/initialed-registers-recommendation-june-11-2010.pdf}. It should be noted, however, that the most recent rulemaking may have broken some new ground in the identification of the class of users and class of use. In particular, ‘noncommercial videos’ were added to the ‘film clip’ exemption, with the Register identifying the class of users as ‘noncommercial, transformative users.’ \textit{Id} at 49. Both this class of uses and class of users seem uncharacteristically broad, and this may signal that the rulemaking process has some chance of growing more flexible in the future.} The rulemaking process has proven to be of little help to the users of copyrighted works; it is even worse for device manufacturers. The rulemaking process applies only to acts of circumvention, and not to devices, nor to their manufacturers, at all. The final way in which the DMCA seeks to mitigate its negative impact is via a series of specific exemptions, set out in sections 1201 (d)-(j).\footnote{17 U.S.C. §§ 1201(d)-(j).} But these exemptions are narrow, confusing, and most are of no help to user innovation.\footnote{See e.g. Pamela Samuelson, \textit{Why The Anti-Circumvention Regulations Need To Be Revised}, 14 BERKELEY TECH. L. J. 519 (1999).}

Given that the built-in safeguards of the DMCA have proven to be of little force, what should the law do to preserve user innovation in the face of the DMCA? An obvious tactic would be to expand and revitalize the rulemaking process, so that it can excuse broader classes of users and types of activities. An even better approach would be to also expand the rulemaking process to cover the manufacture and distribution of tools, in addition to acts of circumvention. However, in many ways the rulemaking process, which was meant to preserve socially valuable uses of copyrighted works, has turned out to be worse than having no process at all. It gives the false appearance of a providing safety valve, and it consumes a significant amount of time and resources, when in reality its narrowness and specificity has done more harm than good in preserving the rights and abilities of users of copyrighted works. In addition, there seems to be little chance that this process will be significantly changed in the years to come.

The best solution, from the point of view of user innovation, would be for the DMCA not to exist at all, but the historical trend in intellectual property law has been towards ever higher levels of protection,\footnote{See e.g. James Boyle, \textit{Cruel, Mean, or Lavish? Economic Analysis, Price Discrimination and Digital Intellectual Property}, 53 VAND. L. REV. 2007, 2010 (2000).} so it is unlikely that there will ever be the political or legislative will to get rid of the DMCA. In the end, the best source of assistance for user innovation would seem to be the courts. Although to date they have been relatively hostile to
both circumvention devices and acts of circumvention, even when they are primarily for fair use purposes, this trend can be reversed if a more robust reading of 1201(c) is applied across all users, uses and devices. Only time will tell if the courts will take this course.

2) Do Technological Controls Interfere with the Development of Feelings of Psychological Ownership?

Before we move on to examine the appropriate legal response to technological developments, we need to address the possible impact that technological controls may have on the reluctance of individuals to bargain over divided rights. We generally take it for granted that when we purchase an asset, we do not need to ask for permission to use it (or to pay for the privilege of its continued use). After all, to paraphrase Samuel Johnson: “[N]obody but a blockhead ever bought except for use.” But perhaps this expectation only exists because we have learned it as a result of experience, and perhaps we might be able to learn a different set of expectations about asset use. In other words, does technology help to shape consumers’ perceptions of what freedoms they should have, by shaping what freedoms they actually do have?

In Chapter 3, we saw that using, controlling and possessing assets leads to the development of feelings of psychological ownership. When legal rights of use are separated from the right of possession, these feelings of psychological ownership give rise to a mismatch between the parameters of legal and psychological ownership, and this mismatch creates a reluctance to bargain over divided rights. However, technological measures can reduce the amount of use and control that users have over assets, so this may mean that users will not develop feelings of psychological ownership, even when they have permanent possession of assets. And this lack of psychological ownership may in turn mean that users will be more likely to ask for permission to make new uses of technologically-controlled assets. If all this is true, then does it justify a greater legal tolerance for governance rules when they are implemented via technology?

127 The original quote by Samuel Johnson is of course: “No man but a blockhead ever wrote, except for money.” ELIZABETH M. KNOWLES, THE OXFORD BOOK OF QUOTATIONS 415 (Oxford 1999). We explored the effect of this expectation on transacting behavior in Chapter 3.
128 See infra Chapter 3, Part 3, Divided Rights vs. Fragmented Rights.
The evidence to date indicates that technological controls are not in themselves enough to interrupt the development of feelings of psychological ownership. Rather than accepting the inherent limitations of technology, users spend a great deal of time and effort seeking to circumvent these measures and bend technology to their will, and neither technology nor lawsuits appear to have changed consumers’ perceptions about the appropriateness of filesharing.129 In other words, technology does indeed interfere with user innovation, by making it more difficult for users to use and modify assets, but it has less of an impact on psychological ownership per se. However, technology is not the only source of potential interference with the development of feelings of psychological ownership – law and social norms are two others.

Like technology, the law can interfere with user innovation in two ways – directly and indirectly. Direct interference occurs when users are less likely to circumvent technological restrictions because they are afraid of getting sued. Indirect interference occurs when the legal rule against circumventing technological restrictions develops into a social norm (thus vitiating feelings of psychological ownership). The first of these scenarios is already occurring - the ‘chilling effect’ of lawsuits has been noted in regards to the DMCA, for example.130 But as with technological controls, the chilling effect of the law appears not to have interfered with feelings of psychological ownership per se. In other words, to date, the legal rule has not developed into a social norm.131

Of course, the law is not the only source of social norms – religion, culture, education, biological hardwiring – are some others.132 Whatever the source, if society adopts

---

129 See e.g. Ken Nicholds, The Free Jammie Movement: Is Making A File Available To Other Users Over A Peer-To-Peer Computer Network Sufficient To Infringe The Copyright Owner's 17 U.S.C. § 106(3) Distribution Right?, 78 FORDHAM L. REV. 983, 991-2 (2009): “The chairman of the RIAA, Mitch Bainwol, claims that the lawsuits have still served an educational purpose: to make it clear to the public at large that downloading is illegal. The EFF concedes that the lawsuits have increased awareness of illegality, but cites several studies demonstrating that the majority of individuals still use filesharing networks even though they believe it is illegal to do so. Without regard to the perception of its legality, one poll found that forty-five percent of Canadians still regard filesharing as something that should be permitted, and another U.K. survey of young people found that almost two-thirds felt as though they should not have to pay for the music that they listen to.”

130 See generally, EFF, supra note 110.

131 Indeed, other researchers have found that: “If anything, digital piracy is the social norm among this segment of consumers.” Rajiv K. Sinha & Naomi Mandel, Preventing Digital Music Piracy: The Carrot or the Stick?, 72 J. MKTG. 1, 13 (2008):

a norm of compliance with use restrictions on assets, and users internalize that norm, then this would presumably interfere with the development of feelings of psychological ownership. There are at least some people who hope that such a norm can be engineered. The movie industry, for example, spends a significant amount of time and money seeking to shape consumer habits. Most of us have seen the ‘anti-piracy’ ads that run in theatres and at the beginning of DVDs; most of us will not be so familiar with the elaborate education programs aimed specifically at children in the classroom. These programs are designed to alter how children think about interacting (and how they actually interact) with IP works and copies.

Advertising campaigns aimed at adults are unlikely to interfere with the feelings of psychological ownership that develop via the possession and use of copies. Although these sorts of public opinion campaigns are sometimes successful (the ‘green movement’ and the anti-smoking movement are two examples), it is arguable that these have been successful because they identify the harm to the audience, not the harm to the advertiser. The educational programs aimed at children have the potential to be somewhat more successful, both because they are integrated within the school curriculum, and because they are aimed at a maturing audience whose ownership expectations are not yet fully formed. However, even in the case of children, the phenomenon of psychological ownership is not infinitely malleable. Children display ownership behavior from a very early age, and this seems to be central to their development of self. While a deep and pervasive set of social norms might be able to redefine the acceptable objects of ownership, it seems unlikely that some

---

133 Some scholars assert that norms must be internalized before they can be considered norms. That debate will not be addressed here. For a discussion of some of the difficulties in defining norms, see e.g. Christine Horne, Sociological Perspectives on the Emergence of Social Norms, in SOCIAL NORMS 4 (Michael Hechter & Karl-Dieter Opp, eds. 2001).
134 See generally Tarleton Gillespie, Characterizing Copyright in the Classroom: The Cultural Work of Anti-Piracy Campaigns, 2 COMMUNICATION, CULTURE, & CRITIQUE 274 (2009).
135 In the case of the green movement, harm to the planet is harm to your own air, land and water. Even drunk driving campaigns work this way, since the ultimate harm to the audience is either personal bodily harm, harm to loved ones, or jail time harming others. It is true that some campaigns succeed in making consumers look beyond their own self-interest, but these campaigns tend to use images that evoke a dramatic emotional response. Green Peace and PETA, in particular, seem to get good mileage out of images that are either cute and cuddly, or bloody and violent. It seems doubtful that copyright owners can draw upon such imagery, so the only avenue left is the threat of jail time. This is a risky strategy when you want consumers to buy your product.
classroom lessons are sufficient to counteract the feelings of psychological ownership that arise from the use and possession of copies of copyrighted works.

For the foreseeable future, it seems unlikely that technological controls, changes in property or intellectual property law, or educational campaigns will interfere with the deeply embedded ownership norms of Western society. However, even if technological measures, laws or social norms can interfere with the development of psychological ownership of assets, we need to ask ourselves if we want this interference to occur. The process of using, coming to know and investing the self into assets has a salutary effect on innovation, and the freedom to use and manipulate assets is key to fostering this process. Psychological ownership has a positive effect on innovation (i.e., psychological ownership increases resource values) in two ways. First, innovations are a by-product of the process of developing psychological ownership – i.e., getting to know the asset and investing the self in the asset. Second, once feelings of psychological ownership have developed, these investments increase (and innovative outcomes are thereby increased). Therefore, if users get accustomed to not interacting freely with assets, this will have a deleterious impact on innovation. In the end, then, the process of developing psychological ownership may itself be an important part of the process of user innovation - and so we may want to be extremely wary of interfering with this process, even when we have the ability to do so. Indeed, the benefits of user innovation indicate that we may want to structure the law so that it fosters the development of feelings of psychological ownership - which is exactly what the common law, via the numerus clausus and its related doctrines, does.

3) The Legal Response to Technology and Business Alternatives to Use Restrictions

In Section 1, we saw that technology makes it more difficult for the law to use its traditional tools (asset type and transaction characteristics) to determine the legal enforceability of use restrictions on assets. In Section 2, we saw that technology alone is insufficient to prevent the development of feelings of psychological ownership in assets,

---

137 Jon Pierce, Tatiana Kostova & Kurt Dirks, The State of Psychological Ownership: Integrating and Extending a Century of Research, 7 REV. GEN. PSYCH. 84, 85 (2003): “There is a diverse literature that suggests that the psychology of possession is well rooted in people socialized by a Western heritage.” (Citations removed.)
even when that technology makes it more difficult for users to use and manipulate the assets that they possess. In this Section, we will briefly address three final issues: the appropriate legal response to technologies that provide notice of use restrictions; the appropriate legal response to technologies that automatically enforce use restrictions (self-enforcing technologies); and business alternatives to use restrictions.

When technology merely provides notice of a use restriction, the courts may be asked to step in and decide whether a violation of that use restriction gives rise to a legal claim. What should the law do in such cases – should the technological means of implementing notice affect the legal enforceability of use restrictions? It should come as no surprise that the user innovation theory of the *numerus clausus* counsels the law to respond to such attempts with great skepticism. Whether the method of providing notice of a use restriction is low-tech (e.g., a label) or high-tech (e.g., a clickwrap contract), courts should look beyond the terms of the notice and instead look at the realities of the transaction to decide whether an asset has been sold, rather than merely leased or licensed. 138 In other words, the courts should continue to refuse to let notice overwhelm the legal categories of contract and property. 139

Of course, this only tells us what the law should do in those cases in which the law is asked to determine the legal implications of a technological control. When technology is self-enforcing, the courts do not need to intervene in order for a use restriction to become operative. However, in these cases, the law (meaning legislatures this time) should at least adopt an attitude of non-interference – that is, lawmakers should let the market decide. 140 Allow sellers to sell assets containing technological protection measures, and allow buyers to circumvent them if they wish to do so. Over time, buyers will either: 1) voice their concerns loud enough that the seller discontinues such efforts; 2) buy a similar good from a seller who does not use such technological controls; or 3) get tired of the time and effort needed to circumvent technological controls and simply put up with them. All of these scenarios have

---

138 *See infra* Chapter 4, Part 2, *IP Copies and Notice: The Realities of the Transaction.*

139 For an overview of how courts treat clickwrap, shrinkwrap and browswrap contracts, see Mark Lemley, *Terms of Use,* 91 MINN. L. REV. 459 (2006). In general, courts are more likely to enforce such licenses when both parties are businesses. This is in full accord with the user innovation theory of the numerus clausus, see Chapter 4, Part 2, *User Innovation and the Mechanisms of Governance.*

140 It should however, continue antitrust oversight, so that when technological controls become standardized across an industry, they have a means to interfere. This is similar to the strategy of mandating the content of contracts when egregious terms become the industry norm. *See e.g.* Friedrich Kessler, *Contracts of Adhesion - Some Thoughts about Freedom of Contract,* 43 COLUM. L. REV. 629 (1943).
played out in the past, and markets have adapted in the face of serious mismatches between seller and buyer preferences about technological controls. Indeed, one might argue that this is what markets are supposed to do, in theory at least. However, what lawmakers should not do is exactly what they did in enacting the DMCA: enact laws that provide statutory backup for technological controls. The threat of legal liability seriously harms a buyer’s ability to signal sellers about the strength and direction of his preferences vis-à-vis technological controls.

For the law, the lessons to be drawn from the user innovation theory of the *numerus clausus* are straightforward. A final question is what lessons the theory has for sellers. In other words, how should sellers respond to the phenomena of user innovation and psychological ownership? We have already set out some advisable business alternatives to technological controls (and other methods of implementing use restrictions) in Chapters 2 and 3, but it is worth summarizing them here. Although these strategies apply to both tangible and intangible assets, their implementation tends to be more sophisticated in the case of intangibles. In response to user innovation, sellers should adopt proactive strategies that give users the freedom to innovate, and which enable sellers to locate, learn from, and adopt user innovations. Such strategies have been thoroughly analyzed by von Hippel and other user innovation researchers, who go even further and recommend that sellers develop and distribute user toolkits that make it easier (rather than harder) for users to innovate.141

In response to psychological ownership, sellers would be wise to give users the ability to manipulate copies in ways that correspond to the parameters of psychological ownership, even if that ability encroaches to some extent upon the right to reproduce an IP-protected work. This is what might be called a ‘copy plus’ model – that is, buyers get somewhat more than they would in a world in which copying was difficult or impossible. After some initial resistance on the part of sellers, this model became the *de facto* model in the game industry for several years, at least until gaming moved from distribution of copies via CDs, to platform gaming and digital/online distribution.

However, if sellers feel that they must implement certain technological controls in order to protect their intellectual property, then they should choose a strategy that does not conflict with psychological ownership. There are three models which sellers can choose from

---

to accomplish this goal. The first strategy would be to implement technological restrictions, but provide countervailing benefits to users. A good example of such a strategy is the ‘unlimited downloads’ policy adopted by several online distributors of digital games.\(^\text{142}\) In this model, users lose the ability to resell the digital copies of the games they download, but they gain the ability to download a purchased game as many times as they want. This means that if they get a new computer, for example, they do not have to purchase a new copy of the game. This model ameliorates feelings of psychological ownership because although the buyer must go through a series of procedures to unlock, register or authenticate his copy of the game, and although he cannot resell his copy of the game, he gets the extra value of portability across computer hardware. Thus, in this model, something is given to the buyer \textit{in exchange for} the intrusions upon his feelings of psychological ownership in regards to the copy.

A second option for sellers is to move towards a hybrid ‘sale/service’ model. This model can be used to replace pure metered use pricing strategies. In this model, a copy of the software product is sold outright, but an ongoing subscription is required for the buyer to receive software updates or ongoing customer service. The buyer’s copy does not suffer ‘lockout’ at the end of the service subscription, but the value of the copy is diminished by the loss of the service package. This model does not conflict with feelings of psychological ownership because the copy gets allocated to the user - i.e., the realm of property; and the service gets allocated to the realm of contract.\(^\text{143}\)

If a seller \textit{is} absolutely intent on implementing technological restrictions that conflict with the development feelings of psychological ownership over copies, then a third strategy is warranted. In this strategy, sellers move to a service model, and keep control of copies

\(^{142}\) See e.g. Direct2Drive’s explanation of ‘download protection’ at http://www.direct2drive.com/staticpage.aspx?topic=about.

\(^{143}\) This is the legal (and psychological) reality, even though software sellers generally insist on referring to the ‘sale’ of a copy of software as a ‘perpetual license.’ See e.g. this explanation of ‘Open Value’ – a volume licensing program that allows companies to purchase Microsoft software along with a package of support services: “Your rights to use the original software licenses themselves (perpetual license rights) are in no way affected by Software Assurance, so you can continue to run the software licenses you originally purchased (and there are no more payments since you completed your 3 year spread payments already); however, by not renewing your Software Assurance, your Software Assurance benefits would expire, meaning you no longer would have: Upgrade Protection, eLearning courses, Office Home Use rights (for Office licenses purchased through Open Value), and all other Software Assurance benefits you were originally eligible for.” http://blogs.msdn.com/b/mssmallbiz/archive/2007/04/03/2018452.aspx.
entirely. Some online games use this model, for example. This strategy does not conflict with feelings of psychological ownership because no copy of the game is ever transferred to the buyer. Instead, the player must ‘visit’ and sign into an online site whenever he wants to play the game. Although most of these ‘service only’ models require regular payments on the parts of users, others rely on periodic ‘lump sum’ payments. These ongoing payments further reinforce the feeling that players are paying for a ‘service,’ rather than for a transfer of property. However, it should be noted that even in this model, sellers have to be wary of the power of psychological ownership. Players tend to develop feelings of ownership, not towards the game itself, but towards their character and virtual ‘possessions’ in the game. Evidence of these feelings of psychological ownership can be found in the importance of ‘virtual property’ in these games, and in the disputes that arise over this ‘virtual property.’ Thus, these games are the most successful when their internal design and administration does not conflict with feelings of psychological ownership.

CONCLUSION

In this Chapter, we examined the implications of the user innovation theory of the *numerus clausus* for intellectual property law, which is one of the most abundant sources of enforceable use restrictions on assets. The nonrival nature of intellectual property assets makes them an ideal target of user innovation. Nonetheless, restrictions on the use, reproduction and adaptation of works are justified by the need to provide incentive for the initial production of a good that is expensive to produce, but cheap to reproduce. We also saw that the law’s differential treatment of land and chattels has been robustly extended to

---

144 See e.g. World of Warcraft FAQ, [http://www.worldofwarcraft.com/info/faq/general.html](http://www.worldofwarcraft.com/info/faq/general.html) (last visited June 15, 2010).

145 See e.g. James Grimmelmann, *Virtual Worlds as Comparative Law*, 49 N.Y.L. SCH. L. REV. 147, 148-49 (2004): “If one had to choose a single canonical feature of multiplayer online games, there would be no contest. Property is invariably among the first features implemented in any game; only the abilities to communicate with other players and to move around the game world are as widespread. Even though it would be technically just as easy to make all virtual items pure public goods, no game has ever gone this route.”

146 “In a landmark decision, Li Hongchen v. Beijing Arctic Ice Technology Development Co., the Beijing Second Intermediate Court (famous for its intellectual property decisions) considered a complaint by a virtual world inhabitant against the holder of the online environment in a dispute over ownership of virtual property. The virtual property of Li Hongchen was taken by a third party when his account was hacked. The lower court determined that Beijing Arctic Ice had an obligation to restore the property to its rightful owner. The court of appeals affirmed the decision, and ordered restoration of the property.” Joshua Fairfield, *Virtual Property*, 85 B.U.L. REV. 1047, 1084-85 (2005) (citations removed).
intellectual property works and copies, but that it breaks down somewhat in the face of technological advances that enable a work to be controlled via direct control of its copies.

Having thus accounted for the current contours of the *numerus clausus* (in Chapter 4 for tangibles, in Chapter 5 for intangibles), we then looked to its future, examining whether advances in technology should result in a relaxation of the doctrine. In contrast to previous theories of the *numerus clausus*, we concluded that a relaxation of the doctrine is *not* warranted, because user innovation and psychological ownership persist even in the face of technological advances that provide notice of (and implementation of) use restrictions on assets. Therefore, both property law and the more specialized regime of intellectual property law would be well-advised to maintain their features which currently serve to limit the enforceability of use restrictions, even in the digital and online environments.
Conclusion

“Think how hard physics would be if particles could think.”

1 Nobel Laureate Murray Gell-Mann, as quoted in DAN ARIELY, PREDICTABLY IRRATIONAL: THE HIDDEN FORCES THAT SHAPE OUR DECISIONS 244 (1st ed. 2008).
This thesis began with a legal puzzle: why does the law refuse to allow property owners to divide their property into as many idiosyncratic bundles as they wish? In Chapter 1 we examined some of the current economic explanations for this puzzle - known as the *numerus clausus* - but found that they did not fit well with the approach of the common law. In particular, these theories focus on the impact that new forms of property have on third parties, and they predict that the law will allow the number of new property forms to increase as the cost of providing notice of these forms decreases. In contrast, the courts focus on the parties to the transaction, and do not let notice determine the boundaries of the *numerus clausus*. We set out to construct a theoretical explanation which would coincide with these features of the law. Our first step was to provide a new definition of the *numerus clausus* – one which equated the principle with the law’s preference for exclusion, and antipathy for governance rules. Rules of exclusion unify rights of use and possession in an asset; governance rules divide these rights.

In Chapter 2, we justified the law’s preference for exclusion in terms of the deleterious impact that governance rules or divided rights have on *user innovation*. Resource values are higher when rights of use and possession are unified because a significant amount of valuable innovation can be revealed only through asset use. Users have obvious advantages over nonusers in both the problem finding and problem solving that emerges through asset use. In addition, because each individual has a unique stock of rival and non-transferable human capital, each possessor interacts differently with an asset than every other possessor of a duplicate asset. Finally, the emergent nature of innovation means that no one (including the innovators themselves) can predict who will innovate, when they will innovate or how they will innovate. This means that the results of user innovation cannot be contracted for in advance (*ex ante*).

In Chapter 3 we examined why individuals will not engage in *ex post* recontracting around governance rules. Most importantly for us, divided rights dramatically decrease the willingness to deal of both the owner of the use right, and the possessor (user) of the asset. As a result, what at first appear to be situations of low transaction costs (e.g., a buyer has permanent possession of an asset, a seller retains use rights, and technology allows low-cost monitoring, enforcement and payment for incremental use), become situations of high transaction costs by virtue of the property rights allocations themselves (e.g., the buyer’s
possession of the asset gives rise to subjective feelings of ownership, which results in attempts to disable the technological mechanisms). Thus, *ex post* rearrangement of divided rights will often not occur, even when a mutually beneficial transaction would otherwise result.

In Chapter 4, we took a closer look at how the user innovation theory of the *numerus clausus* accounts for the law’s approach to divided rights in tangible assets. We saw that the law draws the line between enforceable and unenforceable restrictions on tangible assets (i.e., between governance and exclusion) based on two factors: 1) the type of asset; and 2) the characteristics of the transaction in which the asset is involved. Asset type determines whether there is a need to search out new uses of an asset, while transaction characteristics help determine whether feelings of psychological ownership will develop with respect to an asset.

Where there is a need to search out new uses of an asset, and where users have advantages in undertaking that search, the law has a preference for exclusion. Where there is little need to search out new uses of assets, or where there is a need to coordinate the uses of assets, governance rules are given more latitude. In this way, user innovation accounts for the law’s continued resistance to servitudes on chattels, and its relative tolerance for servitudes on land.

The feelings of psychological ownership that arise from the permanent use and possession of assets account for the law’s preference for exclusion when an asset is permanently transferred to the possessor, and its greater tolerance for governance rules when an asset must be returned to the use right owner. Furthermore, these feelings of psychological ownership are not displaced by mere notice alone, but they can potentially be displaced via the bargaining process that accompanies negotiated contracts. Thus, psychological ownership accounts for the following features of the *numerus clausus*: the law’s intolerance for divided rights (and its tolerance for fragmented rights); the law’s refusal to recognize permanent leases; and the law’s refusal to let notice alone determine the boundaries of the *numerus clausus*.

In Chapter 5 we moved on to examine the implications of user innovation for intellectual property, which is the richest source of enforceable use restrictions on assets. We saw that although intellectual property seems to be anathema to user innovation, intellectual
property law can be justified as a limitation on user innovation because it incentivizes the production of an asset which has public good characteristics. In addition, both copyright and patent law have explicit built-in protections for user innovation, such as fair use and blocking patents. Even more importantly, the inner structure of intellectual property law can itself be understood in terms of the law’s strategy of regulating divided rights according to asset type. The idea/expression dichotomy and the nonobviousness criterion make the intellectual property work more like a piece of land; while the first sale doctrine makes the copy of the intellectual property work more like a chattel. These features enable intellectual property law to regulate works and copies in much the same way that property law regulates land and chattels. Rights in copies and chattels are largely allocated to downstream buyers; while rights to intellectual property works (and, to a lesser extent, land) are largely allocated to upstream owners.

Of course, the effectiveness of this legal strategy breaks down somewhat in the face of technological developments that enable intellectual property owners to control their work via direct control of its copies. Thus, we examined the future of the numerus clausus: how should it respond to technological developments that not only decrease the cost of providing notice of use restrictions, but can sometimes enforce them automatically? Unlike other theories, the user innovation theory does not call for a relaxation of the numerus clausus in the face of technological developments, because user innovation becomes more important with advances in technology. In addition, the psychological ownership that develops from asset use does not decrease in the face of technological controls. Thus, the courts should continue to look at the realities of the transaction, rather than to the mere fact of notice, when evaluating the enforceability of use restrictions. In those cases where technological enforcement of use restrictions is direct, the law should at least refrain from providing legal reinforcement for technological controls. Finally, businesses would do well to choose business models that do not conflict with the development of psychological ownership, and which enable them to share the benefits of user innovation.
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 arcane, too unpredictable, too subject to circumvention by determined parties. These criticisms have been aimed at everything from the legal distinction between ‘license’ and ‘sale’,\(^2\) to the entire law of servitudes.\(^3\) 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’ instincts tell them that things ‘just aren’t right.’ There is some truth to this view – certainly the courts often spend more time asserting the impossibility of novel forms of property, than searching for a solid justification for their instinctive distress.\(^4\) 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 creative lawyering.\(^5\) Other challenges to the *numerus clausus* are indirect, arising from demands for an evolving notion of contractual formation. As technology progresses, so the argument goes, so should our ability to enforce contractual restrictions. In some cases, such legal evolution is asserted to be nothing less than an economic necessity: “[C]onsider the software industry itself. Only a minority of sales take place over the counter, where there are boxes [which have contractual terms printed them] to peruse. … Much software is ordered over the Internet by purchasers who have never seen a box. Increasingly software arrives by wire. There is no box; there is only a stream of electrons, a collection of information that includes data, an application program, instructions… and the terms of sale. … On Zeidenberg’s arguments, these unboxed sales are unfettered by terms--so the seller has made a broad warranty and must pay consequential damages for any shortfalls in performance, two "promises" that if taken seriously would drive prices through the ceiling or return transactions to the horse-and-buggy age.”\(^6\)

---

\(^2\) See *e.g.* Mallinckrodt, Inc. v. Medipart, Inc., 976 F.2d 700, 705 (Fed. Cir. 1992): “That the viability of a restriction should depend on how the transaction is structured was denigrated as "formalistic line drawing" in Continental T.V., Inc. v. GTE Sylvania, Inc., 433 U.S. 36, 57-59, 53 L. Ed. 2d 568, 97 S. Ct. 2549 (1977).”

\(^3\) See *e.g.* C. Haar & L. Leibman, *Property and Law* 703 (1977).


\(^6\) ProCD, Inc. v. Zeidenberg, 86 F.3d 1447, 1451-52 (7th Cir. 1996).
The implication of these criticisms is, of course, that we should abandon the rusty legal fictions of property law entirely, and open the floodgates to the ingenuity of private parties, who simply want to enforce their arrangements in order to implement their business models. But what if we were to look at it another way, as if the vagueness and flexibility were part of the design,\(^7\) 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.\(^8\) 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,\(^9\) is also required for economic activity to occur, and we do want individuals to fulfill their contractual obligations most of the time.\(^10\)

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 can 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

---

\(^7\) 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. The most likely explanation for the existence of the *numerus clausus* is that judges made their decisions based on the same behavioral gut reactions to divisions of use and possession that impact the behavior of consumers and individuals.

\(^8\) Although Henry Smith has made such an attempt, his conclusions are not adopted here. See Henry E. Smith, *Exclusion Versus Governance: Two Strategies for Delineating Property Rights*, 31 J. LEGAL STUDIES S453 (2002).

\(^9\) Even though this ‘coordination’ is often non-negotiated, and thus unilaterally imposed. For this reason, I avoid using the term ‘cooperation’.

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.”

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 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: “Whatever the organizational form of an economy, transaction costs loom large. Under the private-enterprise system they comprise the entire incomes of lawyers, brokers, managers, judges, accountants, policemen, business clerks, secretaries, public servants, and many others.” Indeed, “in the modern world, it would be difficult to find a rich country where transaction costs sum to less than half of nation income.”

But it is difficult, if not impossible, to precisely identify where the ‘unnecessary’ transaction costs (those we should seek to eradicate, like lawyers) end, and the ‘necessary’

---

11 Steven N. S. Cheung, The Transaction Costs Paradigm, 36 ECON. INQUIRY 514, 517 (1998) [Cheung, TCP]. Cheung estimated that under communism, transaction costs in China were 80% of GDP. Id. at 515.
12 There have been many such measures, though very few of them have been comprehensive, due to definitional difficulties. In the most comprehensive attempt, Wallis and North estimated that transaction costs were about 45 per cent of GNP. See John J. Wallis & Douglass C. North, Measuring the Transaction Sector in the American Economy, 1870-1970, in LONG TERM FACTORS IN AMERICAN ECONOMIC GROWTH 95 (Gallman ed., 1986). See also DOUG NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE 28 (Cambridge University Press 1990): “Wallis and North… measuring the size of transaction costs that go through the market (such as costs associated with banking, insurance, finance, wholesale, and retail trade; or, in terms of occupations, with lawyers, accountants, etc.) in the U.S. economy found that more than 45 percent of national income was devoted to transacting and moreover, that this percentage had increased from approximately 25 percent a century earlier.” On the integral nature of transaction costs to economic life see generally Harold Demsetz, The Exchange and Enforcement of Property Rights, 7 J. L. ECON. 11 (1964).
13 STEVEN N.S. CHEUNG, WILL CHINA GO CAPITALIST? 34-5 (Institute of Economic Affairs 1982).
14 Cheung, TCP, supra note 11 at 515.
transaction costs (those that are necessary to a functioning market, like secretaries) begin. This difficulty is illustrated by the controversy in economics over what a transaction cost actually is, and the closely related problem of how to measure them.\textsuperscript{15} Definitions of ‘transaction costs’ have included everything from the costs of transportation (this was an early definition that has now been discredited), to the entire activity of the service sector.\textsuperscript{16} This difficulty arises in economics because it is inherently difficult (if not impossible) to cleanly separate ‘transactions’ from ‘production.’\textsuperscript{17} 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 contractual arrangements.\textsuperscript{18} 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.\textsuperscript{19}

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

\begin{footnotesize}
\begin{itemize}
\item[\textsuperscript{15}] See e.g. Doug Allen, \textit{What are Transaction Costs?}, 14 RES. L. ECON. 1 (1991); and Yoram Barzel, \textit{Transaction Costs: Are They Just Costs?}, 141 J. INST. & THEOR. ECON. 4 (1985).
\item[\textsuperscript{16}] For a clear and exhaustive discussion of these various definitions, and a review of the literature, see Douglas W. Allen, \textit{Transaction Costs, in ENCYCLOPEDIA OF LAW AND ECONOMICS} 893 (Boudewin Bouckeart and Gérriet De Geest eds., Edward Elgar 2000).
\item[\textsuperscript{17}] Id. at 911-912.
\item[\textsuperscript{18}] See e.g. Richard Langlois, \textit{The Vanishing Hand: the Changing Dynamics of Industrial Capitalism}, 12 IND. & CORP. CHANGE 351 (2003).
\end{itemize}
\end{footnotesize}
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 property owner charges money for access to his property.\(^{20}\) 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.\(^{21}\) 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.\(^{22}\) By allowing downstream owners to ignore 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,

---


\(^{21}\) Cheung, TCP, *supra* note 11 at 521-522.

provides us with a rough way to differentiate between ‘necessary’ and ‘unnecessary’ transaction costs. If an individual is simply free riding on the human capital of downstream transferees of an asset - merely adding a layer of costly payments and permissions without adding any independent value\footnote{In this way, the user innovation theory of the \textit{numerus clausus} dovetails to some extent with the labor theory of Locke.} - then there is little reason for the law to enforce such payments and permissions, except where there is actual contractual agreement to them. Usually, the behavior that we might call ‘free riding on human capital’ is associated 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.\footnote{See e.g. Fred S. McChesney, \textit{Rent Extraction and Rent Creation in the Economic Theory of Regulation}, 16 J. LEGAL STUD. 101 (1987).} 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.\footnote{See e.g. Scott Kieff, \textit{IP Transactions: On the Theory & Practice of Commercializing Innovation}, 42 HOU. L. REV. 727, 741 (2005): “In contrast to the real anticommons problem of the postsocialist economy, where an anticommons is correctly described, much of the private value that an IP owner obtains from the IP right to exclude is the ability to openly trade permission for use in exchange for money or other consideration. The economic motivations associated with such "residual claims" are precisely what mitigate anticommons concerns (citations removed).”} 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 make better decisions about resource use 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.26

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.27 This is why possession 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.28 Property law circumvents this insoluble problem by using the device of possession to somewhat arbitrarily separate the exclusionary sheep from the governance goats. Thus, possession has important work to do, not just in determining initial ownership, but whenever an asset is permanently transferred. Indeed, in real and personal property, possession operates in much the same way as authorship operates in copyright law.29 ‘Authorship’ provides a rough (and somewhat fictional30) way to cut off the rights of former creators so that later creators can create without tracing back and compensating all the antecedents of their works. Similarly, possession provides a rough way to cut off the rights of former owners so that later owners can interact with their property free from the proprietary claims of former owners.

And because full use is unified with possession only when the physical transfer of an asset has achieved a certain degree of permanence, the coherence of the property concept is maintained. Possession must walk a fine line in property law, balancing itself between opposing chasms of incoherence. On the one hand, if possession is required at all times for

26 Margaret Radin, Problems for the Theory of Absolute Property Rights, in REINTERPRETING PROPERTY 114 (1993): “the grantors’ welfare in imposing their whims would have to be weighed against this, and whims are hard to weigh.”

27 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 e.g. Pennsylvania Coal Co. v. Mahon, 260 U.S. 393 (1922). In the modern regulatory state, the ability of a property owner to insulate his property from government regulation has been greatly attenuated. Even in some cases, from ‘government’ regulation that is accomplished via a private hand, Kelo v. City of New London, 545 U.S. 469 (2005).

28 Indeed, some scholars equate the right to contract for use with the right to exclusive use, such that there cannot be one without the other. See e.g. Steven N.S. Cheung, The Structure of a Contract and The Theory of a Non-Exclusive Resource, 13 J. L. ECON. 49 (1970).

29 Or invention (as opposed to discovery) in the case of patent law.

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.\textsuperscript{31} 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 is once again destroyed. 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 new owner sells the land, after owning it for only one day, 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 such a resale/reuse royalty. Suppose the legal system adopted an extreme version of the ‘first in time is first in right’ principle: once an initial property owner is identified, ownership never transfers, even after an asset permanently changes hands. In this scheme, all uses made by all future transferees must be paid for, and all uses are subject to the permission of the ‘initial’ owner. Such ‘pay-per-use’ or ‘permission’ property 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 another, but permanently use them for the benefit of another.

What this discussion of resale/reuse royalties reveals is that, for the concept of property to have any meaning at all, we at all times need islands of exclusion in a sea of

governance. 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 downstream use (and 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, because 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.\(^{32}\)

The user innovation theory of the *numerus clausus*, by incorporating data from empirical studies of innovation and insights from such fields as psychology and experimental economics, provides only a general guide to drawing the line between governance and exclusion. But it *does* provide a far more realistic view of how innovation actually occurs, and of how the law can foster or hinder that innovation. 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 general theories to generate testable implications has been repeatedly criticized.\(^{33}\) But perhaps generality is, in the end, a real strength of the law, rather than a 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. In fact, if economics takes itself seriously as a science of human choice, then it should be far more concerned about the untidy reality of actual human behavior, than about its ability to ‘predict’ outcomes based on unrealistic and highly simplified assumptions. Despite economics’ explicit aversion to general theories, and its deification of ‘testable implications,’ the view that realistic assumptions are more important than accurate predictions does find support in a rather surprising place: the work of Ronald Coase. Indeed, economists would do well to take Coase

\(^{32}\) Cheung, TCP, supra note 11 at 515-516.

\(^{33}\) See e.g. OLIVER WILLIAMSON, THE MECHANISMS OF GOVERNANCE 5 (Oxford University Press 1996) [WILLIAMSON].
seriously when he says that we should judge economic theory based on the reasonableness of its assumptions, rather than on its predictive power:

“Most readers will pass over these opening sentences (Putterman omits them when reprinting my article), and others will excuse what they read as a youthful mistake, believing, as so many modern economists do, that we should choose our theories based on the accuracy of their predictions, the realism of the assumptions being utterly irrelevant. I did not believe this in the 1930’s and, as it happens, I still do not. In my Warren Nutter lecture, delivered in 1981, I argue that economists do not, could not, and if they could, should not choose their theories on the basis of the accuracy of their predictions.”34

In fairness, we should recognize that Coase himself had a somewhat optimistic view of bargaining behavior (and of the ability of competition to weed out those who display an irrational unwillingness to deal).35 We should also note that while many mainstream economists remain tied to the grail of mathematical predictability, others have begun to recognize that real human behavior and traditional economic theory are a poor fit.36 Yet other economists have spearheaded the development of entire fields that study how individuals actually make decisions - fields such as behavioral economics. Indeed, from one of these researchers comes a way to at least partially reconcile the opposing values of mathematical predictability and realism. Dan Ariely has posited that human beings are not rational, irrational, nor even boundedly rational,37 but predictably irrational.38 That is to say, human behavior often follows a logic that may be decidedly ‘noneconomic’, but which has its own logic nonetheless.39 Although such ‘regular irregularities’ may be resistant to mathematical modeling, they do give human behavior a sufficient degree of predictability to engage in

36 See e.g. James Buchanan & Yong Yoon, Symmetric Tragedies: Commons and Anti-Commons, 43 J. L. ECON. 1, 12 (2000).
37 That is, they intend to be rational, but they have bounded cognitive capabilities. See e.g. WILLIAMSON, supra note 33 at 42 (noting the concept and its role in transaction cost economics); and HERBERT SIMON, MODELS OF MAN (Wiley 1957) (originator of the concept).
38 See generally Ariely, supra note 1.
39 Id.
what Richard Nelson calls ‘appreciative theorizing.’  

Appreciative theorizing draws upon empirical evidence, but it is verbal rather than mathematical in nature. Indeed, appreciative rather than mathematical theories may be all that we need to construct socially beneficial laws, because “much of our empirical knowledge can be described only qualitatively, with quantitative measures in a supporting but not exhaustive role.”

The user innovation theory of the *numerus clausus* also finds support in the work of another economist: Friedrich von Hayek. One of the central messages of Hayek’s work was that the key to economic success lies in increasing the choice set of the individual. In the end, this is what exclusion is all about: increasing the choice set of those who possess assets. This gives them the freedom of action to explore the full range of asset use, and leads to innovation that cannot be predicted, but which emerges organically out of asset use. So, we need property (exclusion) to help us discover all those things that we otherwise would not know (to uncover all those innovations which would otherwise not emerge). Of course, property, as with all institutions, brings with it its own set of costs, and we oftentimes replace governance with exclusion and get nothing in return except a slightly different distribution of wealth. And we must never lose sight of the fact that exclusion imposes costs on the former owners of assets, who can no longer enforce their governance rules. However, the economic value of exclusion arises from the fact that the long-term costs of governance will be greater, overall, than the short-term costs of exclusion. Indeed, what we might call ‘the failure rate’ of exclusion is a general cost of increasing the choice set of individuals:

“If we knew how freedom would be used, the case for it would largely disappear. We shall never get the benefits of freedom, never obtain those unforeseeable new developments for which it provides the opportunity, if it is not also granted where the uses made of it by some do not seem desirable. It is therefore no argument against individual freedom that it is frequently abused. Freedom necessarily means that many things will be done which we do not like. Our faith in freedom does not rest on the foreseeable results in particular circumstances but on the

---

41 Id. at 5.
belief that it will, on balance, release more forces for the good than for the bad.”

We could just as easily read this quote to refer to ‘property’ instead of ‘freedom’. Quite literally, if we knew why and when we needed property (a rule of exclusion), if we knew who would generate the most valuable new uses of resources, and what those new uses actually would be, we would not need property at all. We could simply devise a complex web of payments and permissions (governance rules), a perfect scheme in which every person would give his best and get his due, and society would benefit as a result. But we cannot do this, and so we will never know what the ‘right’ mix of governance and exclusion is. What we do know, however, is that innovation is inherently unpredictable because it is the result of an emergent process. We also know that human capital plays a critical role in promoting economic growth over time, and that each person’s stock of human capital is rival, nontransferable and unique. In addition, we know that when rights of use and possession are divided, human beings will develop feelings of psychological ownership that go beyond the boundaries of legal ownership, and those feelings hinder the process of recontracting over divided rights. And so, what we do know demands that we keep a close eye on the relative amount of governance and exclusion in the economy, no matter how technologically or organizationally sophisticated we may become. Better to make our slow but sure way via horse and buggy, than to put the cart before the horse.

---

Bibliography

Case Law

Austerberry v. Oldham (1885) 29 Ch. D 750.
Central P. R. Co. v. Alameda County, 284 U.S. 463 (1932).
Everex Sys. v. Cadtrak Corp. (In re CFLC, Inc.), 89 F.3d 673 (9th Cir. 1996).
In Re Ellenborough Park [1956] Ch. 131.
Jarvis v. K2, 486 F.3d 526 (9th Cir. 2007).


LEE v. A.R.T. Co., 125 F.3d 580 (7th Cir. 1997).

Lexmark v. Static Control Components, 387 F.3d 522 (6th Cir. 1994).


Miles Medical Co. v. John D. Park & Sons Co., 220 U.S. 373 (1911).


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

Pennsylvanica Coal Co. v. Mahon, 260 U.S. 393 (1922).


ProCD, Inc. v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996).


Revlon v. Logisticon, No. 70533 (Super. Ct., Santa Clara County, filed Oct. 22, 1990);


Robinson v. Campbell, 16 U.S. 212 (1818).
Sony Computer Entm't Am., Inc. v. GameMasters, 87 F. Supp. 2d 976 (N. D. Cal. 1999).
Sperry & Hutchinson Co v Fenster, 219 F 755 (ED NY 1915).
Stowe v. Thomas, 23 F. Cas. 201 (C.C.E.D. Pa. 1853).
TransWorld Airlines, Inc v American Coupon Exchange, Inc, 913 F2d 676 (9th Cir 1990).
Tri-Continental v. Tropical Marine Enterprises, 265 F2d 619 (5th Cir 1959).
United States v. Causby, 328 U.S. 256 (1946).
Universal City Studios Inc. v. Corley, 273 F.3d 429 (2d Cir. 2001).

Secondary Sources

ABA SECTION OF ANTITRUST LAW, HANDBOOK ON THE ANTITRUST ASPECTS OF STANDARDS SETTING 74 (2004).


APPROVES: WEBSTER’S QUOTATIONS, FACTS AND PHRASES (ICON Group International 2008).


**Yoram Barzel, Economic Analysis of Property Rights** (2d ed. 1997).


Donald Campbell, *Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes*, 67 PSYCHOLOGICAL REVIEW 380 (1960).

W.B. Cannon, *The Role of Chance in Discovery*, 50 SCIENTIFIC MONTHLY 204-09 (1940).
Canon EOS-1D Mark II N Firmware License Agreement, available at http://web.canon.jp/Imaging/eos1dm2n/eos1dmark2nfirmware-e.html (last visited June 11, 2010).


OPEN INNOVATION: RESEARCHING A NEW PARADIGM (Henry William Chesbrough, Wim Vanhaverbeke, Joel West, eds., 2006).


STEVEN N.S. CHEUNG, WILL CHINA GO CAPITALIST? (Institute of Economic Affairs 1982).


Ronald Coase, The Nature of the Firm, 4 ECONOMICA 386 (1937) [Coase].


Harold Demsetz, Toward a Theory of Property Rights, 57 AMER. ECON. REV. 347 (1967).


CHARLES ELLWOOD, CULTURAL EVOLUTION: A STUDY OF SOCIAL ORIGINS AND DEVELOPMENT (1927).


Paul Goldstein, *Copyright’s Highway: From Gutenberg to the Celestial Jukebox* (Rev. ed. 1994).


OLIVER HART, FIRMS, CONTRACTS AND FINANCIAL STRUCTURE (Oxford University Press 1995).


ELHANAN HELPMAN, GENERAL PURPOSE TECHNOLOGIES AND ECONOMIC GROWTH (MIT 1998).


V. DANIEL HUNT, ALBERT PUGLIA AND MIKE PUGLIA, RFID – A GUIDE TO RADIO FREQUENCY IDENTIFICATION (Wiley 2007).

MARTIN HUSZ, HUMAN CAPITAL, ENDOGENOUS GROWTH, AND GOVERNMENT POLICY (P. Lang 1998).


David Jacobson, Low Tech and R&D Spending: Problems in Policy for Innovation (April 2005) (Manuscript, on file with author.)


MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES (1987).


Mark Lemley, Terms of Use, 91 Minn. L. Rev. 459 (2006).


DUNCAN MATTHEWS, GLOBALIZING INTELLECTUAL PROPERTY RIGHTS: THE TRIPS AGREEMENT (Routledge 2002).


Paul Muller & Julien Penin, *Why Do Firms Disclose Knowledge and How Does it Matter?*, in *Innovation, Industrial Dynamics and Structural Transformation* 149 (Uwe Cantner & Franco Malerba, eds., 2006).

**RICHARD NELSON, TECHNOLOGY, INSTITUTIONS AND ECONOMIC GROWTH** (2005).

**RICHARD R. NELSON & SIDNEY G. WINTER, AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE** (Belknap Press 1982).


**OECD, MEASURING WHAT PEOPLE KNOW** (1996).

**PAUL ORMEROD, WHY MOST THINGS FAIL** (2005).


Margaret Radin, Humans, Computers, and Binding Commitment, 75 Ind. L.J. 1125 (2000).


David A. Rice, Licensing the Use of Computer Program Copies and the Copyright Act First Sale Doctrine, 30 JURIMETRICS J. 157 (1990).


Paul Romer, Endogenous Technological Change, 98 J. POL. ECONOMY S71 (1990)


Pamela Samuelson, Why The Anti-Circumvention Regulations Need To Be Revised, 14 BERKELEY TECH. L. J. 519 (1999).


**JIM TAYLOR**, *EVERYTHING YOU EVER WANTED TO KNOW ABOUT DVD* (2004).


FREDERICK WASSER, VENI, VIDI, VIDEO (University of Texas Press 2001).


JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET—AND HOW TO STOP IT (Yale University Press 2009).

Statutes


Government Documents


Recommendation of the Register of Copyrights in RM 2008-8 June 11, 2010, pages 63-64

Scottish Law Commission, Discussion Paper No. 112, Discussion Paper on Conversion of

Scottish Law Commission No. 204, Report on Conversion of Long Leases. December 2006,

Sustainable Development Working Group, Sustainable Development Indicators, Working