



**GEORGE
MASON**
UNIVERSITY

School of Law

The Rules and Standards of Patentable Subject-Matter

**Tun-Jen Chiang,
George Mason University School of Law**

***Wisconsin Law Review*, Forthcoming**

**George Mason University Law and Economics
Research Paper Series**

10-42

This paper can be downloaded without charge from the Social Science
Research Network at http://ssrn.com/abstract_id=1670871

The Rules and Standards of Patentable Subject-Matter

Tun-Jen Chiang[†]

Two arguments are commonly made against restricting patentable subject-matter. The first is that such restrictions are over-inclusive. If an invention is “new, useful, and non-obvious,” critics ask, why should it be denied patent incentives because it falls into some “wrong” category? The second criticism is that patentable subject-matter doctrine is difficult to administer, with no coherent principle to explain the case law in the area.

When viewed from a rules versus standards perspective, these arguments contradict each other. Over-inclusiveness is an attribute of rules. Difficulty of administration, vagueness, and inconsistent application are attributes of standards. A doctrine cannot be too rigid and too fuzzy at the same time.

This Article refutes both criticisms of subject-matter doctrine. The insight is that patentable subject-matter doctrine comes in two distinct types. The first is a rule-like categorical exclusion. The second is a standard-like scope limitation, which does not pose problems of over-inclusiveness, while sharing the same heightened administrative cost as other aspects of individualized examination. Since individualized examination is the only alternative to subject-matter restriction, flexible scope limits should not concern critics of subject-matter restriction.

The remaining concern is the over-inclusiveness of categorical exclusion. This Article argues that categorical exclusions can be justified if they create corresponding administrative cost savings that outweigh the over-inclusiveness cost, and this cost-benefit balance is an empirical question. For example, if 99% of business method patents are socially detrimental, it is likely better to categorically reject all business method patents using a simple rule, and accepting the loss of 1% of meritorious patents.

[†] Assistant Professor of Law, George Mason University School of Law. Thanks to John Duffy, Michael Risch, and participants at the Works in Progress IP Colloquium for comments on an earlier version of this article. Research funding from the Law & Economics Center is also gratefully acknowledged.

INTRODUCTION	2
I. THE EXCESSIVE MONOPOLY THEORY AND ITS PROBLEMS	6
A. The Excessive Monopoly Cost Rationale	6
B. The Over-Inclusiveness Problem	8
II. CATEGORICAL EXCLUSIONS, INDIVIDUALIZED EXCLUSIONS, AND SCOPE LIMITATIONS	12
A. Background on the Specification and Claims.....	12
B. Exclusions Versus Scope Limitations.....	16
1. Scope limitation and the cost-benefit balancing problem. ...	16
2. Exclusions.	22
C. Categorical Versus Individualized Exclusion.....	25
1. Categorical exclusions.	25
2. Individual exclusions.	27
D. Subject-Matter Doctrine Contains Both Doctrines of Scope Limitation and Doctrines of Categorical Exclusion.....	29
1. The clarity and over-inclusiveness of subject-matter rules.	30
2. The judicial struggle to define abstract ideas.....	34
E. The Importance of Distinguishing Categorical Exclusion from Scope Limitation.....	40
1. Confused doctrine: <i>In re Bilski</i>	40
2. Confused academic discussion.....	43
III. AN ADMINISTRATIVE COST THEORY OF CATEGORICAL EXCLUSIONS	45
A. Rules Versus Standards	46
B. Layers of Patent Scrutiny	49
C. Principles for Categorical Exclusion.....	53
1. Cost-benefit balance.....	53
2. Ease of administration and definition.	57
CONCLUSION	60

INTRODUCTION

Two arguments are commonly made against restricting patentable subject-matter. The first argument is that such restrictions are rigid and over-inclusive. We shouldn't exclude *all* business methods from patentability, for example, if the problem is only a subset of business method inventions are trivial or obvious, and we can use more finely-tuned doctrines such as obviousness to prevent patents on that subset from issuing.¹ The second argument is that the doctrines on

¹ Michael Risch, *Everything is Patentable*, 75 Tenn. L. Rev. 591, 658 (2008) (“the PTO and courts should focus on answering specific questions about how

patentable subject-matter are fuzzy and difficult to apply.² In this view, the problem is that the category of “business method” is too malleable and uncertain.³

When viewed from the perspective of the classic rules versus standards debate, these two arguments contradict each other. Rigidity and over-inclusiveness are features of rules. Conversely, fuzziness and inconsistent application are features of standards. To the extent the definition of a “business method” is rigid, it cannot be fuzzy at the same time; and to the extent that the definition of a “business method” is too malleable, that flexibility should prevent over-inclusiveness.

The contradiction is explained if we separate out *individual doctrines* of subject-matter restriction. Section 101 of the Patent Act broadly defines patentable subject-matter as “anything under the sun that is made by man,”⁴ but courts have imposed several judge-made restrictions under the rubric of this section. Various decisions have stated that the following are not patentable subject matter:

- 1) Laws of nature or scientific principles
- 2) Abstract ideas
- 3) Natural phenomena or processes
- 4) Natural products

Other restrictions, such as traditional prohibitions on patenting business methods and software, have been overruled in recent decades,⁵ though courts have shown renewed interest in enforcing subject-matter limits.⁶

to best apply rigorous standards of novelty, nonobviousness, utility, and specification with a scalpel rather than simply eliminating broad swaths of innovation with a machete”); Kristin Osenga, *Ants, Elephant Guns, and Statutory Subject Matter*, 39 Ariz. St. L.J. 1087, 1115–18 (2007) (arguing that subject-matter exclusion is an overbroad “elephant gun”).

² Risch, *supra* at note 1, at 607 (criticizing “unclear and undefined subject matter rules”).

³ See John R. Allison, Mark A. Lemley & Joshua Walker, *Extreme Value or Trolls on Top? The Characteristics of the Most-Litigated Patents*, 158 U. Pa. L. Rev. 1, 7 (2009) (“Business-method patents are notoriously difficult to define, with possible definitions varying greatly in scope.”).

⁴ *Diamond. v. Chakrabarty*, 447 U.S. 303, 309 (1980).

⁵ *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*, 149 F.3d 1368, 1375 (Fed. Cir. 1998) (“We take this opportunity to lay this ill-conceived [business method] exception to rest.”).

⁶ See, e.g., *CyberSource Corp. v. Retail Decisions, Inc.*, 620 F.Supp.2d 1068, 1081 (N.D. Cal. 2009) (“*Bilski*’s holding suggests a perilous future for most

Generally, courts and commentators lump all subject-matter doctrines (also known as section 101 doctrines) together for purposes of analysis.⁷ This is a mistake because, as this Article outlines, different restrictions have very different positions on the rules versus standards continuum. Some subject matter doctrines function as bright-line rules of *categorical exclusion*. The prototypical example is the exclusion of natural products. The exclusion is quite clear: there is almost no ambiguity about what constitutes a natural product. At the same time, the categorical exclusion can be over-inclusive. Even if a naturally occurring herb is found to cure cancer, it cannot be patented.

On the other end of the spectrum, some subject-matter doctrines function as flexible *scope limitations*. The prototypical example is the prohibition on patenting abstract ideas and scientific principles. This “abstract idea” doctrine is not capable of bright-line rules: all patents cover “ideas.”⁸ Furthermore, the patented ideas exist at multiple levels of abstraction,⁹ so the Wright Brothers’ invention can be described very concretely as a single barely-working glider made of wood, or very abstractly as “flying machines” that covers all later developments such as the space shuttle and jet fighters.¹⁰ Because there is no bright-line rule on how much abstraction is *too* abstract, the “abstract idea” doctrine is an indeterminate standard.

Once subject-matter doctrines have been sorted between categorical rules and flexible standards, a more intelligent analysis can be made. The scope limitations of section 101 require little further discussion. Like other standards governing patent scope, they are unpredictable for litigants, expensive to administer for courts and the Patent Office, but can be precisely tailored to prevent excessive monopolies. Despite

business method patents.”); *see also* *Bilski v. Kappos*, 130 S. Ct. 3218, 3231 (2010) (“[N]othing in today's opinion should be read as endorsing interpretations of § 101 that the Court of Appeals for the Federal Circuit has used in the past. . . . In disapproving an exclusive machine-or-transformation test, we by no means foreclose the Federal Circuit's development of other limiting criteria.”).

⁷ *See In re Bilski*, 545 F.3d 943, 952 n.5 (Fed. Cir. 2008) (*en banc*), *aff'd*, 130 S. Ct. 3218 (2010) (lumping all the restrictions together under the heading of unpatentable “fundamental principles”).

⁸ *See Mazer v. Stein*, 347 U.S. 201, 217 (1954) (noting that copyrights protect only expression while patents protect the idea itself).

⁹ Tun-Jen Chiang, *The Levels of Abstraction Problem in Patent Law*, at 26 online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1434465 (last visited February 19, 2010) (describing the multiple levels of abstraction that any patent will contain).

¹⁰ *Id.*

all the administrative costs of this type of standard, critics of section 101 doctrine can point to no alternative, since the alternative to a flexible standard are the categorical rules that section 101 doctrine is more conventionally associated with.

What is needed then, is a justification for the categorical exclusions that are implemented under the rubric of section 101. Here, critics of subject-matter restriction have often ignored a key factor: the low administrative costs of categorical rules. The categorical aspects of subject-matter doctrine—exclusions of natural products and natural phenomena (and, previously, living things), as well as the statutory exclusions of human clones and nuclear weapons—have proven clear and easy to administer, and thus requiring little litigation. This is made apparent by the comparison to the “abstract ideas” doctrine, which has been indeterminate. Ironically, the fact that the “abstract ideas” doctrine is indeterminate and thus more frequently litigated has, through the availability heuristic,¹¹ led courts and commentators to conflate its problems with all of section 101 doctrine.

The administrative cost advantage of categorical subject-matter exclusions is important when considered within the context of the finite resources of the U.S. Patent and Trademark Office (“PTO”), a classically overworked and understaffed government agency.¹² To individually examine every patent application, given the PTO’s resource constraints, is to examine every application equally *badly*.¹³

This Article contributes to the literature on patentable subject-matter by dividing section 101 into rules and standards; and provides utilitarian justifications for each type of subject-matter restriction. Part I provides a background on the utilitarian justification for patents, and the criticisms of subject-matter restriction made under this theory. Part II then separates out section 101 doctrine into two types: categorical exclusions that function as rules, and scope limitations that function as standards. Part III turns to an analysis of the role of categorical exclusions, and argues that categorical exclusions can be

¹¹ See generally Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 4 *Cognitive Psychology* 207 (1973) (describing the availability heuristic).

¹² See Adam B. Jaffe & Josh Lerner, *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It* 133-73 (2004).

¹³ See Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 *Berkeley Tech. L.J.* 577, 596 (1999) (“patent applications should be subject to differing levels of scrutiny depending on how much social cost they entail”).

justified under utilitarian theory as a method of reducing the administrative costs of the patent system.

I. THE EXCESSIVE MONOPOLY THEORY AND ITS PROBLEMS

A. *The Excessive Monopoly Cost Rationale*

A variety of rationales are commonly advanced to justify restrictions on patent eligibility. The legalistic argument is that the Constitution limits patentable subject-matter to the “useful Arts,” an eighteenth century term that is largely synonymous with “technology” in modern usage.¹⁴ The moral argument, usually raised by gene patents and other patents related to living things, is that some types of inventions are either themselves immoral, or their propertization violates anti-commodification norms.¹⁵ The utilitarian argument is that patents, contrary to their stated purpose, often end up reducing social welfare by imposing excessive monopoly costs.¹⁶

Far above the others, the excessive monopoly cost argument is the one that receives the most scholarly and judicial attention.¹⁷ Justice Breyer’s dissent from the dismissal of certiorari in *Laboratory Corp. of America Holdings v. Metabolite Laboratories, Inc.* (“*Labcorp*”), a case about a patent on diagnostic tests for vitamin deficiency, provides a good example of the excessive monopoly cost argument:¹⁸

¹⁴ Karl B. Lutz, *Patents and Science: A Clarification of the Patent Clause of the U.S. Constitution*, 18 Geo. Wash. L. Rev. 50, 54 (1949) (“The term ‘useful arts’ as used in the Constitution ... is best represented in modern language by the word ‘technology.’”). Contrary to popular conception, the Constitution does not authorize the grant of patents for the advancement of science, since “science” as used in the Constitution denoted learning in general, and was the domain of copyrights. *Id.* at 51-52.

¹⁵ Esther Slater McDonald, Note, *Patenting Human Life and the Rebirth of the Thirteenth Amendment*, 78 Notre Dame L. Rev. 1359 (2003) (arguing that patents covering a human embryo violate the Thirteenth Amendment); Margo A. Bagley, *Patent First, Ask Questions Later: Morality and Biotechnology in Patent Law*, 45 Wm. & Mary L. Rev. 469, 503-04 (2003) (arguing that a patent may represent a “badge or incident of slavery”); Rebecca Dresser, *Ethical and Legal Issues in Patenting New Animal Life*, 28 Jurimetrics J. 399, 410-24 (1988) (discussing arguments against patenting animals).

¹⁶ See, e.g., Harry First, *Controlling the Intellectual Property Grab: Protect Innovation, Not Innovators*, 38 Rutgers L.J. 365, 376 (2007) (arguing for “parsimony” in granting intellectual property rights).

¹⁷ See, e.g., *Bilski*, 130 S. Ct. at 3253-57 (Stevens, J., concurring) (“The primary concern is that patents on business methods may prohibit a wide swath of legitimate competition and innovation.”).

¹⁸ 548 U.S. 124 (2006).

The relevant principle of law excludes from patent protection laws of nature, natural phenomena, and abstract ideas. . . . [T]he reason for the exclusion is that sometimes too much patent protection can impede rather than promote the Progress of Science and useful Arts [P]atents do not only encourage research by providing monetary incentives for invention. Sometimes their presence can discourage research by impeding the free exchange of information, for example by forcing researchers to avoid the use of potentially patented ideas . . . and by raising the costs of using the patented information, sometimes prohibitively so.¹⁹

The excessive monopoly cost theory thus calls for patentable subject-matter to be determined by balancing the incentive benefits of offering patent protection in a field with the monopoly cost of the resulting patents. Patents should be denied if the cost-benefit balance is unfavorable. Rochelle Dreyfuss has made this case explicitly with respect to patents on business methods:

The bottom line is this. The costs of business method patents are very high. The benefits, at least the traditional benefits, are low. The ratio is terrible. The case for patents on business methods is simply not there, at least not in general.²⁰

Similar arguments are made with respect to other fields such as biotechnology and basic science.²¹ The argument, framed in more

¹⁹ *Id.* at 126-27 (Breyer, J., dissenting from dismissal of certiorari) (internal quotations and alterations omitted).

²⁰ Rochelle Cooper Dreyfuss, *Are Business Method Patents Bad for Business?*, 16 Santa Clara Computer & High Tech. L.J. 263, 276-77 (2000).

²¹ *See, e.g.*, Peter Yun-Hyoung Lee, *Inverting the Logic of Scientific Discovery: Applying Common law Patentable Subject Matter Doctrine to Constrain Patents on Biotechnology Research Tools*, 19 Harv. J.L. & Tech. 79, 101 (2005) (arguing that “courts and prospective patent holders should not extend the patent system to enable a ‘monopoly of knowledge’ over basic intellectual assets that could ‘block off whole areas of scientific development, without compensating benefit to the public.”); Jay Dratler, Jr., *Does Lord Darcy Yet Live? The Case Against Software and Business-Method Patents*, 43 Santa Clara L. Rev. 823, 833-36 (2003) (arguing that patents on business methods allow monopolization of entire lines of business); Julia Alpert Gladstone, *Why Patenting Information Technology and Business Methods Is Not Sound Policy: Lessons from History and Prophecies for the Future*, 25 Hamline L. Rev. 217, 230-232 (2002) (arguing that business method patents do not incentivize innovation); Malla Pollack, *The Multiple*

general terms, is that patent incentives are unnecessary in many fields, and the cost of monopoly outweighs the marginal benefits provided by patents in those areas.²²

B. *The Over-Inclusiveness Problem*

The problem with the excessive monopoly cost theory when applied to patentable subject-matter restrictions is one of over-inclusiveness in the exclusion. For example, one can readily agree with Professor Dreyfuss's contention that the benefits of *most* business method patents are low, and that *some* business method patents create significant monopoly costs.²³ Yet it does not follow from the above that *all* business methods should be categorically excluded from patentability. Instead, supporters of business method patents (and opponents of subject-matter restrictions more generally) argue that the solution is to more finely separate the wheat from the chaff. That is, we should reject *individually* the bad business method patents; not eliminate the entire category of all business methods from

Unconstitutionality of Business Method Patents: Common Sense, Congressional Consideration, and Constitutional History, 28 Rutgers Computer & Tech. L.J. 61, 61-62 (2002) (“Since business methods are ‘useful’ when they directly earn revenue, they are inherently unlikely to be under-produced due to market failure.”).

²² See David S. Olson, *Taking the Utilitarian Basis for Patent Law Seriously: The Case for Restricting Patentable Subject-Matter*, 82 Temple L. Rev. 181, 221-32 (2009); Alan Devlin & Neel Sukhatme, *Self-Realizing Inventions and the Utilitarian Foundation of Patent Law*, 51 Wm. & Mary L. Rev. 897 (2009) (arguing that “self-realizing” inventions should be excluded under section 101); see also Peter S. Menell, *A Method for Reforming the Patent System*, 13 Mich. Telecomm. & Tech. L. Rev. 487, 501-02 (2007) (arguing for a system where different fields of technology would receive different treatment based on cost-benefit balancing); Michael A. Carrier, *Unraveling the Patent-Antitrust Paradox*, 150 U. Pa. L. Rev. 761, 826-27 (arguing that “patents are not necessary for innovation in many industries”).

²³ This must be considered in the context that the overwhelming majority of issued patents are worthless and thus generate neither benefit for their owners nor monopoly cost to the public. See Kimberly A. Moore, *Worthless Patents*, 20 Berkeley Tech. L.J. 1521 (2005)

patentability. John Duffy is a leading proponent of this view,²⁴ though it is frequently articulated by other scholars as well.²⁵

The way to achieve the utilitarian goal of the patent system, from this point of view, is to individually scrutinize patent applications under the criteria of novelty, usefulness, and non-obviousness.²⁶ This argument becomes especially tempting because obviousness is a malleable concept.²⁷ Thus, one can define non-obvious inventions as only those that would not have been invented or disclosed “but for the inducement of a patent.”²⁸ Under such a definition—which collapses the test into the basic utilitarian goal of patents—a non-obvious patent is by definition a “good” patent in the broader normative sense.²⁹ If the invention would not exist without the patent, then granting the patent is cost-free: all upside and no downside.³⁰ Any additional restrictions

²⁴ John F. Duffy, *Rules and Standards on the Forefront of Patentability*, 51 Wm. & Mary L. Rev. 609, 622-23 (2009) (“it should be a rare situation in which an entire class of patents complies with the nonobviousness requirement and yet still somehow discourages or impedes the development and spread of useful knowledge”).

²⁵ Risch, *supra* at note 1, at 606-09; Osenga, *supra* at note 1, at 1115-18; *see also Bilski v. Kappos*, 130 S. Ct. 3218, 3229 (2010) (“Finally, even if a particular business method [satisfies § 101], that does not mean that the application claiming that method should be granted. In order to receive patent protection, any claimed invention must be novel, § 102, nonobvious, § 103, and fully and particularly described, § 112.”).

²⁶ Duffy, *supra* at note 24, at at 623 (“there are alternatives to restricting patentable subject matter rules, including imposing a more demanding nonobviousness requirement”).

²⁷ *Harries v. Air King Prods. Co.*, 183 F.2d 158, 162 (2d Cir. 1950) (Hand, J.) (describing non-obviousness as “as fugitive, impalpable, wayward, and vague a phantom as exists in the whole paraphernalia of legal concepts.”).

²⁸ *Graham v. John Deere Co.*, 383 U.S. 1, 11 (1966) (giving this standard as the theoretical policy goal of the obviousness doctrine); *see Tun-Jen Chiang, A Cost-Benefit Approach to Patent Obviousness*, 82 St. John’s L. Rev. 39 (2007) (formulating an economic standard for obviousness determinations).

²⁹ The other utilitarian restriction is that some inventions we never want to induce at all, such as illegal narcotics or methods of tax evasion. These can similarly be addressed by a very flexible definition of what constitutes a “useful” invention, by flexibly defining usefulness to include only socially beneficial inventions.

³⁰ John W. Schlicher, *Patent Law: Legal and Economic Principles* § 1.04[6][a] (2002) (“A ‘monopoly’ . . . is not bad, if the prospect of that ‘monopoly’ was what induced the . . . invention. The only product ‘monopolized’ and sold at too low quantities is a product that would have been sold in even lower quantity, zero, if there were no ‘monopoly.’”).

on patentability—including all subject-matter restrictions—become meaningless and redundant at best and welfare-reducing at worst.³¹

Of course, one straightforward response is that current obviousness law does not really implement the “but for the inducement of the patent” standard in practice.³² For example, serendipitous or “Eureka!” inventions such as the Post-It Note required virtually no research and thus would probably have been invented even without patents.³³ Despite this, such “flash of genius” inventions are legally considered the height of non-obviousness,³⁴ while inventions requiring large investments and painstaking-but-mundane research are given much less protection.³⁵ But saying that the non-obviousness standard is imperfect simply invites the answer that the non-obviousness standard should be improved.³⁶ It does very little to recommend wholesale exclusion as the alternative. In other words, resort to the even blunter—and therefore even *more* over- and under-inclusive—instrument of categorical subject-matter exclusion does not logically follow from pointing out the over- and under-inclusiveness of current non-obviousness doctrine.

³¹ See Richard S. Gruner, *Intangible Inventions: Patentable Subject-Matter for an Information Age*, 35 Loy L.A. L. Rev. 355, 368-69 (2002) (“recognizing that a category of innovations constitutes patentable subject matter simply secures the opportunity for a few innovations within that category that are new, nonobvious, and subject to timely patent applications to qualify for patent protections and rewards”); Jared Earl Grusd, *Internet Business Methods: What Role Does and Should Patent Law Play?*, 4 Va. J.L. & Tech. 9 (1998) (arguing that courts can prevent issuance of non-meritorious business method patents by “manipulating” the non-obviousness and enablement requirements).

³² Devlin & Sukhatme, *supra* at note 22, at 919.

³³ Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 Va. L. Rev. 1575, 1581 n.13 (2003); see U.S. Patent No. 5,153,041 (filed Oct. 10, 1990) (Post-It Note).

³⁴ *Cuno Eng’g Corp. v. Automatic Devices Corp.*, 314 U.S. 84, 91 (1941) (defining non-obviousness as a “flash of creative genius”).

³⁵ See, e.g., *PharmaStem Therapeutics, Inc. v. ViaCell, Inc.*, 491 F.3d 1342, 1363 (Fed. Cir. 2007) (invalidating patent on a widely-acclaimed medical breakthrough achieved after painstaking experimentation because the researchers had “merely used routine research methods”); see also *Feist Publications, Inc. v. Rural Telephone Serv. Co.*, 499 U.S. 340, 352-53 (1991) (“sweat of the brow” not sufficient to warrant copyright protection).

³⁶ Duffy, *supra* at note 24, at 623 (“Moreover, if such a class of patents can be identified, there are alternatives to restricting patentable subject matter rules, including imposing a more demanding nonobviousness requirement.”); Risch, *supra* at note 1, at 606-09; Osenga, *supra* at note 1, at 1115-18.

The more pointed objection is that individualized examination entails significant administrative costs; and the more flexible this inquiry is, the costlier it becomes. Under a very open-ended obviousness inquiry—whether the invention would exist “but for the inducement of the patent”—a fact-finder would need to find out what would have happened if patent incentives were never provided. This counter-factual is almost impossible to construct accurately.³⁷ Even a more restrained test of obviousness—such as asking whether a person in the field would consider the advances over prior knowledge to be significant—still requires collecting significant amounts of information about the opinions of those skilled in a field, the content of what was previously known, and a weighing of the significance of the differences.³⁸ Because the public domain (known as “prior art”) is extremely vast, this exercise is very costly. Subject-matter restrictions, which are imposed at the threshold and searching the prior art, avoids these costs and ought to be administratively cheaper.³⁹ The problem here, however, is that subject-matter restrictions have shown no sign of being cheap or easy to implement. Instead, “patentable subject matter has long been considered one of the most challenging doctrinal areas of patent law.”⁴⁰ To argue that employing patentable subject-matter restrictions will lower administrative cost is therefore a claim that requires some explanation.

This Article explains how subject-matter limits can and do provide lower administrative costs, but only as to particular kinds of subject-matter restriction. As discussed in Part II.D, what is conventionally known as “subject-matter doctrine” really encompasses two distinct types of doctrine: rule-like categorical exclusions and standard-like scope limitations. Categorical exclusions can offer lower administrative cost, though at the cost of over-inclusiveness error. Standard-like scope limitations are difficult and expensive to administer, but allow precise tailoring to utilitarian cost-benefit balancing that is no different from, or even superior to, current non-obviousness doctrine. By separating out these distinct types of subject-matter doctrines, the apparent contradiction in section 101 appearing

³⁷ See Chiang, *supra* at note 28, at 75 (noting that the “key variables of benefit, cost, and timing of independent invention are difficult to predict before the patent is issued”).

³⁸ See *Graham*, 383 U.S. at 17-18 (laying down a four-factor test).

³⁹ Olson, *supra* at note 22, at 203-04.

⁴⁰ Duffy, *supra* at note 24, at 616-38 (discussing the history of subject-matter tests).

to be both overly rigid⁴¹ *and* unacceptably fuzzy⁴² at the same time is resolved.

II. CATEGORICAL EXCLUSIONS, INDIVIDUALIZED EXCLUSIONS, AND SCOPE LIMITATIONS

Broadly speaking, patent law (not limited to subject-matter doctrine) has three types of restrictions on patent rights. Two of these may be termed *exclusions*. That is, patent law allows no patent at all to issue when an exclusion applies to the patentee or his invention. The first type of exclusion is categorical, based on the category of the invention or its inventor rather than the individual merits of either. For example, under the Patent Act of 1793, only U.S. citizens could obtain patents.⁴³ Thus, all non-citizens were categorically excluded from patentability, regardless of the merits of their inventions. The second type of exclusion is individualized, based on the merits of the invention, and is reflected in current law by the requirements of novelty, usefulness, and non-obviousness.⁴⁴

The other limit on patent rights is not an exclusion but a *scope limitation*, which arise after entitlement to a patent is established. Obviously, no patent can be of unlimited scope, and patent protection is limited in both duration and breadth.⁴⁵ Examples of scope limitations include that patents protect only against infringements occurring during the patent term and within the United States.⁴⁶ The most important limit, however, is that patents cover only the “invention” of the patentee.⁴⁷ A background on the patent document is useful in analyzing these restrictions on patent rights.

A. *Background on the Specification and Claims*

A United States patent application is a complex document, but at its heart is the “specification,” which contains a written description of the

⁴¹ Risch, *supra* at note 1, at 658 (“the PTO and courts should focus on answering specific questions about how to best apply rigorous standards of novelty, nonobviousness, utility, and specification with a scalpel rather than simply eliminating broad swaths of innovation with a machete”).

⁴² *Id.* at 647 (arguing that the current judicial tests for subject-matter are uncertain).

⁴³ Patent Act of 1793, Ch. 11, § 1, 1 Stat. 318, 318.

⁴⁴ *See* 35 U.S.C. § 102 (novelty), § 103 (obviousness). The requirement of utility is often traced to section 101. *Brenner v. Manson*, 383 U.S. 519, 528-29 (1966). Utility, however, is entirely distinct from subject-matter.

⁴⁵ 35 U.S.C. § 154 (2006) (defining the grant and term of a patent).

⁴⁶ 35 U.S.C. § 271(a) (2006) (defining infringement).

⁴⁷ *Id.*

patentee's invention.⁴⁸ This description is highly detailed: it must contain sufficient detail to allow any ordinary person in the field to “make and use” the invention.⁴⁹ As the words “make and use” imply, the invention being described is a physical thing or process: since intangible ideas cannot be made, nor even “used” in an observable manner. The thing or process—an *embodiment*—is constructively placed in the public domain by the patent disclosure. If words are inadequate to provide a full disclosure, an actual deposit of a physical sample may be used instead.⁵⁰

Once the specification has been filed, it must be examined by the PTO to see whether a patent should issue.⁵¹ Initially, the question is whether the embodiment fits within the statutory subject-matter categories of process, machine, manufacture, or composition of matter. Because the basic structure of the description requirement already requires that the embodiment be physical—it must either be actually deposited or constructively described so that it can be “made and used”—meeting this literal statutory test is not a serious concern. The second layer of subject-matter doctrine considers whether the embodiment falls within one of the subject-matter exclusions, such as for natural products and processes, atomic weapons, and human clones. Even here, few inventions fall within a proscribed category.

Second, the patentability determination asks whether the embodiment is new, useful, and non-obvious. The first two conditions are far less important than the last. Novelty asks whether the invention is strictly identical to the prior art;⁵² and few people will slavishly replicate what has come before. Similarly, few people will seek to patent a useless invention, since a monopoly on something useless would be worthless, while pursuing the patent costs money, making the pursuit uneconomical.⁵³ The key question for patentability is thus usually whether the invention is non-obvious: *i.e.* whether

⁴⁸ 35 U.S.C. § 112 (2006).

⁴⁹ *Id.*

⁵⁰ *Enzo Biochem, Inc. v. Gen-Probe Inc.*, 323 F.3d 956, 965 (Fed. Cir. 2002); *In re Wands*, 858 F.2d 731, 735 (Fed. Cir. 1988).

⁵¹ “Patentability” is often misused to refer exclusively to refer to subject-matter exclusions. However, the patent statute includes other conditions such as novelty and non-obviousness in the concept of patentability. 35 U.S.C. §§ 101-103 (2006) (titled “conditions of patentability”).

⁵² *Trintec Indus., Inc. v. Top-U.S.A. Corp.*, 295 F.3d 1292, 1296 (Fed. Cir. 2002) (discussing “the strict identity required of the test for novelty”).

⁵³ Robert P. Merges & John F. Duffy, *Patent Law and Policy: Cases and Materials* 207 (4th ed. 2007) (“[A] truly useless invention should be worthless, so who would go through the expense of patenting it?”).

others in the same field would have found it reasonably easy to construct the same embodiment based modifying on similar (but not strictly identical) devices.⁵⁴ If the embodiment passes all these tests, it is deemed patentable.

Historically, the patent then issued with a specification describing the patentee's embodiment, and questions regarding the patent's scope would be resolved later in litigation.⁵⁵ Specifically, patent scope would be determined by juries who were charged with answering whether an accused infringing product and the patentee's original embodiment were, "substantially, in their principles and mode of operation, like."⁵⁶ This regime, however, was considered unsatisfactory, as it left both patentees and potential competitors uncertain about the scope of the patent monopoly. "Substantial similarity," being already an extremely vague test, was made more so when applied by unpredictable juries.⁵⁷

The solution that was adopted over the course of several decades was the evolution of "claims." A claim is a one sentence description, by the patentee, of the scope of his desired monopoly,⁵⁸ reflecting an inventive idea or principle.⁵⁹ Somewhat confusingly, the scope of the claimed monopoly—defined by an idea or principle—is also frequently called the "invention."⁶⁰ But these two concepts of "invention" must be distinguished carefully. A patentee's specification contains one (or very few) embodiment-inventions. From that single embodiment, a patentee can claim a broad idea.⁶¹ For example, the Wright Brothers

⁵⁴ See generally *Non-Obviousness: The Ultimate Condition of Patentability* (John F. Witherspoon ed., 1980).

⁵⁵ See Merges & Duffy, *supra* at note 53, at 782 (noting also that "a patent did not even have to include claims prior to 1836").

⁵⁶ *Odiorne v Winkley*, 18 Fed. Cas. 581, 582 (C.C.D. Mass. 1814).

⁵⁷ See John F. Duffy, *The Festo Decision and the Return of the Supreme Court to the Bar of Patents*, 2002 Sup. Ct. Rev. 273, 309-10.

⁵⁸ *In re Vogel*, 422 F.2d 438, 442 (C.C.P.A. 1970) ("A claim is a group of words defining only the boundary of the patent monopoly.").

⁵⁹ Merges & Duffy, *supra* at note 53, at 27 ("For purposes of the patent law, an invention is only the concept or principle that is articulated in the patent claim.").

⁶⁰ *Id.*; see also *Gill v. United States*, 160 U.S. 426, 434 (1896) ("In every case the idea conceived is the invention."); Giles S. Rich, *The Relation Between Patent Practices and the Anti-Monopoly Laws—Part II*, 24 J. Pat. Office Soc'y 159 (1942), reprinted in, 14 Fed. Cir. B.J. 21, 29 (2004) (the invention is "an incorporeal, intangible abstraction")

⁶¹ Robin C. Feldman, *The Inventor's Contribution*, 2005 UCLA J.L. & Tech. 6, ¶ 60 (2005) ("A patent holder need only identify a single use and a single embodiment for the product to receive rights to a wide range of embodiments and all uses.").

constructed, and their patent specification described, a wooden airplane that barely flew; but their claimed invention was *all* airplanes, including later-developed aluminum jets.⁶²

The distinction between inventions as embodiments and inventions as claimed monopoly scope leads to a corollary, which is the distinction between exclusions of the embodiment and limits on monopoly scope. An exclusion focuses on the specification embodiment;⁶³ and if the specification embodiment fails one of patent law’s criteria, no patent will issue.⁶⁴ In contrast, if the embodiment itself is patentable, the remaining question is the appropriate claim scope. Because a patentee may file as many claims as he likes (and amend them without limit),⁶⁵ a patentee can always eventually obtain some degree of claim scope—as long as there is a patentable embodiment.⁶⁶ Defects in the specification embodiment cannot be remedied by paperwork in the patent office;⁶⁷ excessive claim scope can always be.⁶⁸

With this background, the remainder of this Part provides a more detailed description of the three types of restrictions on patent rights: categorical exclusion, individualized exclusion, and scope limitation. Part II.B differentiates scope limitations from exclusions; while Part II.C differentiates between categorical and individualized exclusions. Part II.D then shows how this taxonomy is useful in analyzing subject-matter doctrine.

⁶² U.S. Patent No. 821,393 (filed Mar. 23, 1903); see Tun-Jen Chiang, *Fixing Patent Boundaries*, 108 Mich. L. Rev. 523, 537 (2010).

⁶³ See, e.g., Howard T. Markey, *Why Not the Statute?*, 65 J. Pat. Off Soc’y 331, 333 (1983) (“Only embodiments of an idea, i.e. an invention, may be patented.”); see also *Seymour v. Osborne*, 78 U.S. 516, 552 (1870) (“in order to constitute an invention, the party must have proceeded so far as to have reduced his idea to practice, and embodied it in some distinct form”).

⁶⁴ Of course, it remains technically true that “it is impossible . . . to ever finally reject a patent application.” Mark A. Lemley & Kimberly A. Moore, *Ending Abuse of Patent Continuations*, 84 B.U. L. Rev. 63, 64 (2004). The patent applicant always has the legal right to file more claims. The point, however, is that these new claims will not have any success, if the fundamental defect is that the specification embodiment is excluded, whereas an overbreadth problem can be cured with a narrower claim.

⁶⁵ *Tafas v. Dudas*, 541 F. Supp. 2d 805, 816 (E.D. Va. 2008).

⁶⁶ See *infra* text accompanying notes 77–80 for a more detailed explanation. In simplified terms, the reason is that if the embodiment is patentable, a patentee would at a minimum be able to obtain claim scope covering only exact replication of the embodiment.

⁶⁷ See 35 U.S.C. § 132 (2006) (prohibiting the addition of “new matter” to the specification).

⁶⁸ Lemley & Moore, *supra* at note 64, at 64.

B. Exclusions Versus Scope Limitations

1. Scope limitation and the cost-benefit balancing problem.

The scope of a patent is defined by its length and breadth. Although patent length is rigidly set by statute,⁶⁹ the breadth of a patent is one of the most flexible policy levers in the patent system that judges control. To see why, it is important to understand that a simplistic equation of the patentee's monopoly to what the patentee creates—what I call the unitary invention theory—simply does not work.⁷⁰

It is often tempting to merge the two conceptions of “invention” by holding that the claimed monopoly may extend *only* to the embodiments taught by the specification. Indeed, numerous cases and other authorities stand for this proposition.⁷¹ The idea of *quid pro quo*—and thus equating contribution to monopoly scope—is fundamental to patent law.⁷² It also has the helpful effect of removing judicial discretion over patent scope.

At the same time, however, actually enforcing such a rule is untenable. The reason is that in that the rule would permit easy evasion of patents, by allowing an infringer to make immaterial alternations that were not (and, if the variant is developed later, could not be) taught by the patentee.⁷³ For example, a patentee who created a wooden table would not be able to cover a later plastic table using the same principles of geometry, because he could not teach how to make plastic at the time of his conception and patent filing.⁷⁴ Equating

⁶⁹ 35 U.S.C. § 154 (2006) (term of patent).

⁷⁰ Chiang, *supra* at note 9, at 23-24 (describing the unitary invention theory and its problems).

⁷¹ *Martek Biosciences Corp. v. Nutrinova, Inc.*, 579 F.3d 1363, 1378 (Fed. Cir. 2009) (“specification must teach those skilled in the art how to make and use the full scope of the claimed invention” (quoting *In re Wright*, 999 F.2d 1557, 1561 (Fed. Cir. 1993)); Robert P. Merges & Richard Nelson, *On the Complex Economics of Claim Scope*, 90 Colum. L. Rev. 839, 845 (1990) (“Under section 112, the disclosure must be sufficient to enable someone skilled in the art to make and use all the embodiments of the invention claimed in the patent.”).

⁷² Giles S. Rich, *Principles of Patentability*, 28 Geo. Wash. L. Rev. 393, 402 (1960) (“That is one of the beauties of the patent system. The reward is measured automatically by the popularity of the contribution.”).

⁷³ Chiang, *supra* at note 9, at 17-20.

⁷⁴ Nor does it work to ask whether a person of skill in the art could use the later-developed technology at the *time of infringement*. By definition, at the time of infringement, a person of skill in the art *has* used the later-developed technology—the infringer. Thus, changing the timing of the inquiry renders the test toothless. See *infra* text accompanying notes 188–190.

patent scope with the embodiments taught by the specification would thus eviscerate patent scope and diminish patent incentives. Not surprisingly in light of this analysis, numerous other cases establish that patent scope can exceed what is taught by the specification.⁷⁵

The upshot is that every patentee teaches one or a few embodiments, but then claims a monopoly defined by a broader idea covering untaught embodiments.⁷⁶ The fuzziness of patent scope lies in the fact that every embodiment can be translated to an infinite array of ideas, each representing a different level of abstraction, and each corresponding to a different degree of patent scope.⁷⁷

Consider the example of a radiation machine that uses X radiation to cure AIDS. This is an embodiment. But how should we define the inventive idea? Some obvious candidates are as follows:

1. Curing AIDS using X radiation.
2. Curing AIDS using radiation.
3. Curing AIDS.

Choosing a level of abstraction to define patent protection is as important as it is difficult.⁷⁸ A patent that covers *all* cures for AIDS is much broader than one that excludes only *radiation machines* curing AIDS; which in turn is superior to one that excluded only using *X radiation* to cure AIDS. The following diagram demonstrates how,

⁷⁵ See, e.g., *Morley Sewing Mach. Co. v. Lancaster*, 129 U.S. 263, 273 (1889) (“all subsequent machines which employ substantially the same means to accomplish the same result are infringements, although the subsequent machine may contain improvements”); *Tilghman v. Proctor*, 102 U.S. 707, 728 (1880) (requiring only “a description of the process and of *one* practical mode in which it may be applied.” (emphasis added)); *Spectra-Physics, Inc. v. Coherent, Inc.*, 827 F.2d 1524, 1533 (Fed. Cir. 1987) (claim not invalid even if it “reads on another embodiment of the invention which is inadequately disclosed”).

⁷⁶ See Emerson Stringham, *Double Patenting* 209 (1933) (“In patent law there is no possibility of clear thinking until it is understood that an ‘invention’ *as protected* . . . is an abstraction, an idea of means.” (emphasis added)); see also *In re Vogel*, 422 F.2d 438, 442 (C.C.P.A. 1970) (“A claim is a group of words defining only the boundary of the patent monopoly. It may not describe any physical thing and indeed may encompass physical things not yet dreamed of.”).

⁷⁷ Chiang, *supra* at note 9, at 26.

⁷⁸ See *Royal Typewriter Co. v. Remington Rand, Inc.*, 168 F.2d 691, 693-94 (2d Cir. 1948) (Hand, J.) (permissible abstraction is “always a question of degree, and courts have differed, and always will differ, as to the allowable latitude in a given instance”).

from a core embodiment, potential patent protection radiates outwards. Note that a subsequently developed competitor product—an AIDS pill—is covered only by a high degree of abstraction. In this way, abstraction and scope are directly correlated.⁷⁹

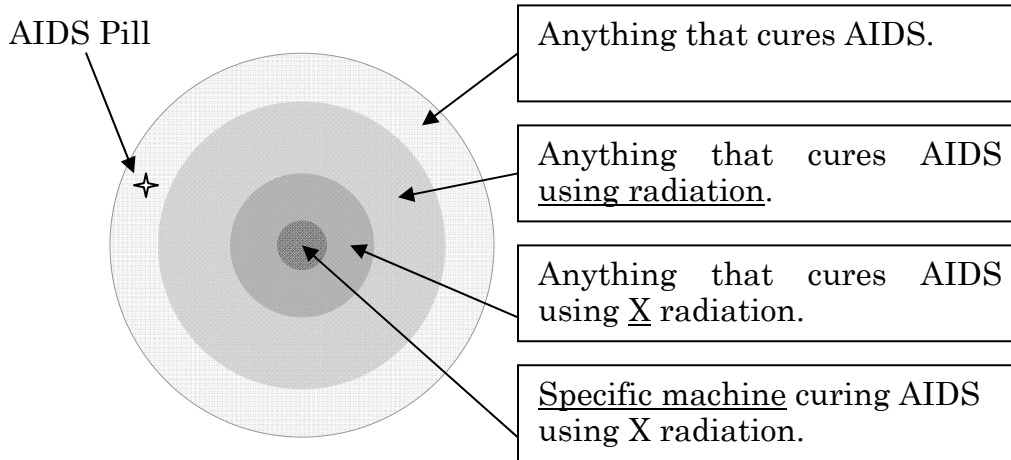


Figure 1: Levels of Abstraction as Patent Scope

The fact that every patented invention exists as a spectrum of levels of abstraction creates a doctrinal problem and a corresponding opportunity. The doctrinal problem is that it is almost impossible to create bright-line rules on what constitutes the “correct” level of abstraction for patent scope.⁸⁰ The problem is reflected in the case law surrounding the doctrine of “enablement,” which is the primary patent law doctrine governing scope.⁸¹ One line of enablement cases confine patent scope to the embodiments taught (or “enabled”) by the

⁷⁹ See William M. Landes & Richard A. Posner, *The Economic Structure of Intellectual Property Law* 323 (2003) (noting that breadth and abstraction are directly correlated); see also *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121 (2d Cir. 1930) (Hand, J.) (outlining a similar levels of abstraction problem in the context of copyright).

⁸⁰ *Royal Typewriter*, 168 F.2d at 693-94.

⁸¹ See Kevin Emerson Collins, *Enabling After-Arising Technology*, 34 J. Corp. L. 1083, 1088 (2009) (discussing “doctrinal chaos” in enablement); Bernard H. Chao, *Rethinking Enablement in the Predictable Arts*, 2009 Stan. Tech. L. Rev. 3, ¶¶ 50–52 (discussing conflicts in doctrine); see also Jeffrey Lefstin, *The Formal Structure of Patent Law and the Limits of Enablement*, 23 Berkeley Tech. L.J. 1141, 1189-90 (2008) (arguing that enablement doctrine is inherently inadequate to govern patent scope).

specification,⁸² and thus eviscerates patent scope to nothingness. Another line of enablement cases disavow any such strict limit and seem to permit infinite patent scope through ever-increasing abstraction.⁸³ These cases are impossible to reconcile with each other and leave no coherent rule for lower courts to apply.

One solution is to allow patentees to define the level of abstraction, by defining the invention through a claim.⁸⁴ However, allowing patentees to define the level of abstraction will predictably result in patents of unlimited scope, so judges retain ultimate control through determining whether the patentee's claim is valid.⁸⁵

The bottom line is that courts retain control and have no real doctrine to bind them. This doctrinal confusion, however, provides a corresponding pragmatic opportunity. For in practice “courts exercise discretion and oscillate between the full-scope and single-embodiment doctrines to achieve the desired outcome. When they feel like the inventor has overreached, they . . . invalidate the claim,” but they invoke the contrary doctrine to uphold what are perceived as deserving claims.⁸⁶ Occasionally, courts have even acknowledged this pragmatic

⁸² See, e.g., *Incandescent Lamp Patent*, 159 U.S. 465, 472-74 (1895) (failure to enable later-developed bamboo filament); *Auto. Techs. Int'l, Inc. v. BMW of N. Am., Inc.*, 501 F.3d 1274, 1284 (Fed. Cir. 2007) (failure to enable later-developed electronic sensor); *Liebel-Flarsheim Co. v. Medrad, Inc.*, 481 F.3d 1371, 1378-79 (Fed. Cir. 2007) (failure to enable later-developed jacketless injector).

⁸³ See, e.g., *Innogenetics, N.V. v. Abbott Labs.*, 512 F.3d 1363, 1371-72 (Fed. Cir. 2008) (“Our case law allows for after-arising technology to be captured within the literal scope of valid claims that are drafted broadly enough.”); *Chiron Corp. v. Genentech, Inc.*, 363 F.3d 1247, 1254 (Fed. Cir. 2004) (“The law does not expect an applicant to disclose knowledge invented or developed after the filing date.”); *Spectra-Physics, Inc. v. Coherent, Inc.*, 827 F.2d 1524, 1533 (Fed. Cir. 1987) (claim not invalid even if it “reads on another embodiment of the invention which is inadequately disclosed”).

⁸⁴ In fact, modern patent claims evolved precisely as a patentee attempt to define the invention at high levels of abstraction. See Duffy, *supra* at note 57, at 309-10 (noting that claims were developed by patentees to “define[] their invention in broad conceptual terms and assert their rights to the invention in those terms”).

⁸⁵ *Hybritech Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384 (Fed. Cir. 1986) (holding that enablement is a legal question); cf. *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 391 (1996) (holding that claim interpretation is “purely legal”). As a practical matter, judges often creatively “interpret” rather than outright invalidate claims to achieve a desired level of scope.

⁸⁶ Collins, *supra* at note 81, at 1088-89.

use of their discretion over patent scope.⁸⁷ The opportunity provided by messy doctrine is that courts have freedom to engage in economic balancing to match reward to the contribution of the inventor.⁸⁸

The benefit of judicial fine-tuning *ex post* comes at significant cost. The first cost is the difficulty of administration: courts must spend time and effort on this exercise, as must PTO examiners. The other cost is the resulting error. In theory, the goal of scope delineation is to balance incentives and monopoly costs; but courts lack the information to make this decision correctly. Any mistakes in this regard, however, creates error costs: grant too little scope, and the incentive for innovation suffers; grant too broad scope, and consumers suffer.⁸⁹ Because lack the administrative resources to collect perfect information, a large degree of error is inevitable.⁹⁰

Despite the high degree of error and expense of administration, it is difficult to construct a better regime to resolve the level of abstraction problem. The only alternative to *ex post* judicial line-drawing is *ex ante* legislative line-drawing.⁹¹ A close point of comparison is the patent term, which like patent breadth involves a balance of incentive benefits and monopoly costs,⁹² but is rigidly set by statute. But although patent term thus acquires the characteristics of a rule and is

⁸⁷ The best example is *Eibel Process Co. v. Minn. & Ont. Paper Co.*, 261 U.S. 45, 63 (1923):

In administering the patent law, the court first looks into the art to find what the real merit of the alleged discovery or invention is and whether it has advanced the art substantially. If it has done so, then the court is liberal in its construction of the patent, to secure to the inventor the reward he deserves. If what he has done works only a slight step forward, and that which he says is a discovery is on the borderline between mere mechanical change and real invention, then his patent, if sustained, will be given a narrow scope, and infringement will be found only in approximate copies of the new device.

⁸⁸ *See id.*; see also Chiang, *supra* at note 9, at 43.

⁸⁹ *See generally* Merges & Nelson, *supra* at note 71; Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 Tex. L. Rev. 989 (1997).

⁹⁰ *Cf. Nash v. CBS, Inc.*, 899 F.2d 1537, 1541 (7th Cir. 1990) (“Neither Congress nor the courts has the information that would allow it to determine [optimal copyright scope]. Both institutions must muddle through.”).

⁹¹ *See id.*

⁹² *See* Richard Gilbert & Carl Shapiro, *Optimal Patent Length and Breadth*, 21 RAND J. Econ. 106 (1990) (noting that patent length and breadth can be traded-off against each other).

easy to administer, it has *even higher* error costs than the messy patent breadth inquiry.

This might seem counter-intuitive, in that few judges are likely to make mistakes in calculating patent term. But *mistakes* in applying rule are not what I mean by “error cost” in this rules-versus-standards context. Instead, by “error cost,” I mean the degree of deviation from *the underlying policy goal*—in the case of the patent system, optimal balancing of incentives and monopoly cost. Mistakes in applying a rule will usually incur error cost, in so far as a rule generally serves its underlying goal. For example, wrongfully acquitting a driver of speeding will usually, but not always, reduce road safety, in that speed limits generally serve the underlying goal of road safety. But imperfections of the rule itself also incur error cost, even when the rule is strictly applied. Not every speeding driver is a menace to traffic safety; and not every driver under the speed limit is a safe one. The over- and under-inclusiveness of the rule itself is a source of error cost. And because nobody thinks that a 20-year term is optimal across-the-board for all patents,⁹³ the rigid statutory term incurs high error costs despite its easy administration. In contrast, a court that is at least *trying* to reach the optimal patent breadth is likely to get closer to the mark, even if it lacks perfect information and thus will still make errors. But it achieves greater perfection at the price of greater administrative cost.

Because the question of scope is inherently complex, it becomes expensive no matter how the cost is allocated.⁹⁴ Done by rule, as in patent length, it creates extremely high error costs. Done by a imperfectly-administered standard, as in patent breadth, it creates high administrative costs and still rather high error costs. If we really insisted on a perfect administration of the standard—to reduce error

⁹³ The literature on optimal patent term is enormous, and there is little consensus beyond the high-level goal of maximizing welfare. *See generally id.* at 107; Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 Harv. L. Rev. 1813, 1823-29 (1984); William D. Nordhaus, *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change* 76 (1969). But because every patent has different effects on welfare, a constant 20 year term is unlikely to be optimal for all but a miniscule number of patents.

⁹⁴ Even complex policy choices can be made through a system of rules, but they become a system of highly complex rules akin to the tax code. *See* Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 Duke L.J. 557, 586-96 (1992) (discussing the interaction between rules versus standards and complexity versus simplicity). The alternative is to abdicate the responsibility and enforce an artificially simple rule, which would be cheap to administer but then leads to absurd results (*i.e.* high error costs).

costs to zero and have optimal scope for every patent—then administrative costs will become almost infinite. None of these options are good.

What remains is to avoid the scope delineation problem when possible. Patent law achieves this through doctrines of exclusion. When an invention is excluded from patentability altogether, courts never have to reach the difficult scope inquiry. In this way, the patent system conserves administrative resources.⁹⁵

2. Exclusions.

In contrast to doctrines of scope that concern ideas and abstraction, doctrines of exclusion focus on whether the patentee’s core embodiment can be patented. The conditions of patentability, including as novelty, usefulness, and non-obviousness, are basic examples of exclusions.⁹⁶ If the patentee’s embodiment in the specification is itself already known, then no patent will issue on it no matter how narrowly it is claimed. This can be represented diagrammatically:

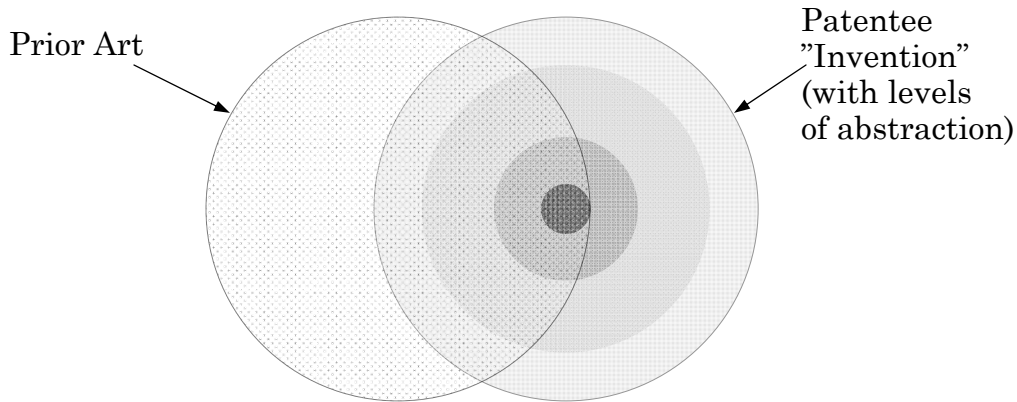


Figure 2: Novelty Exclusion

As seen in Figure 2, when the prior art encompasses the core embodiment, no patent may issue at all. Even if patent scope were to be narrowed only to the specific embodiment described (the dark dot in the middle), it would still overlap with the prior art and thus be invalid.

A similar analysis occurs with the obviousness doctrine. The refinement is simply that instead of the prior art including only

⁹⁵ For discussion on this “layered” approach as a method of conserving administrative resources, see *infra* Part III.B.

⁹⁶ See 35 U.S.C. §§ 101–103 (2006) (labeled “conditions of patentability”).

embodiments that were in the public domain, the prior art now includes those embodiments *plus* their obvious variants.

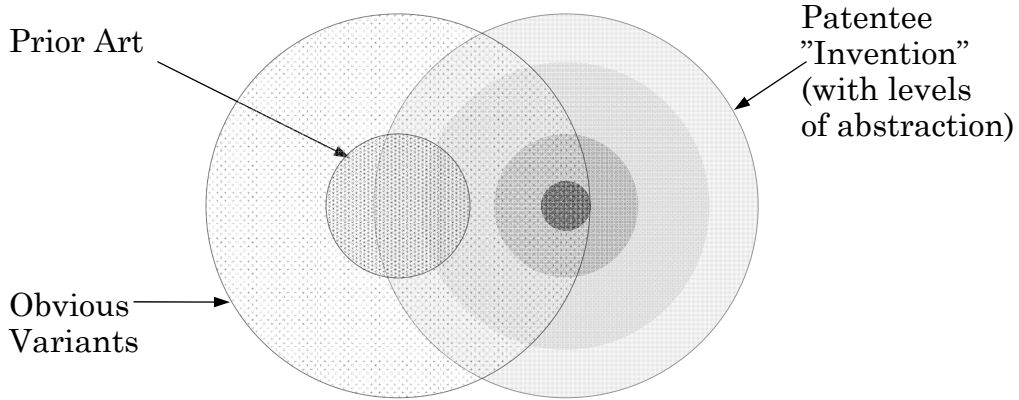


Figure 3: Obviousness Exclusion

This distinction between exclusion and scope limitation is likely to be highly counter-intuitive to many practicing patent lawyers. This is because, in practice, novelty and non-obviousness are primarily invoked to limit patent scope, not to reject patents completely.⁹⁷ Figure 4 provides an illustration of this phenomenon using the example of a fountain pen embodiment.

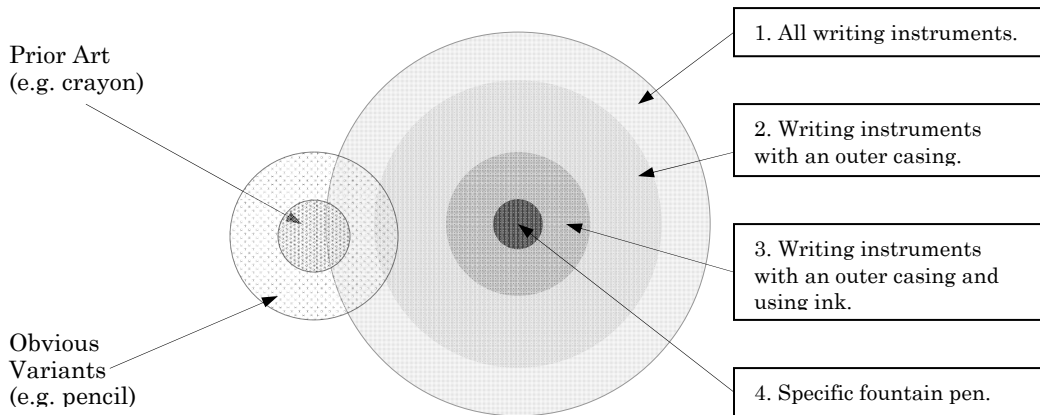


Figure 4: Obviousness Limiting Scope

⁹⁷ See *Wilson Sporting Goods Co. v. David Geoffrey & Assoc.*, 904 F.2d 677, 684 (Fed. Cir. 1990) (limiting scope by prior art).

In this example, the patentee who describes a fountain pen embodiment in the specification has four claims, ranging from the highest level of abstraction (all writing instruments) to the lowest (a single fountain pen). The examiner will reject the broadest claim as non-novel because crayons, which already exist, are also writing instruments. The examiner will further reject the claim to writing instruments with an outer casing as obvious, because pencils are an obvious variant of pre-existing crayons. By rejecting the broad claims, the level of abstraction is lowered.

The fact that novelty and non-obviousness are often used to limit scope, however, does not change their role as doctrines of exclusion rather than scope delineation; for novelty and non-obviousness are not helpful in ultimately resolving the scope question. In the example above, there are still infinite levels of abstraction between “a *single* specific fountain pen,” and “*all* writing instruments with an outer casing and using ink,” including, for example, coverage of ballpoint pens, quill pens, and inkjet printer cartridges. A PTO examiner or court must still ultimately draw the line on the levels of abstraction problem, and excluding the broadest claims are of little help in this *conceptual* problem. The subset of an infinite set (of levels of abstraction) is still an infinite set.⁹⁸

Nor can we avoid the scope delineation problem by permitting the patentee to cover everything that is new and non-obvious. The problem here is that such a rule will predictably result in patents that are obviously too broad in the sense of creating excessive monopoly cost relative to the contribution of the patentee. In the example above, the inventor of a fountain pen embodiment will be permitted to cover future ballpoint pens and inkjet printers. To take this to the logical extreme, the inventor of any new thing, no matter how insignificant, will be able to claim “all new and non-obvious compositions of matter that are made in the next 20 years.”⁹⁹ Plainly, even after novelty and

⁹⁸ I emphasize here that I am speaking of the problem in conceptual terms. Practically speaking, PTO examiners who bear the burden of proving a patentee claim is of undue scope under indeterminate standards are loath to attempt this. See *In re Wright*, 999 F.2d 1557, 1561-62 (Fed. Cir. 1993) (PTO bears initial burden of proof). Thus, examiners strongly prefer to use the more determinate rules of novelty to limit scope, with the result that most novel and non-obvious claims will be allowed unless the scope is wildly excessive.

⁹⁹ The PTO rejects such “omnibus claims” under its style requirements. See *Ex parte Fressola*, 27 U.S.P.Q.2d (BNA) 1608, 1612 (B.P.A.I. 1993). That it is only a stylistic problem means a patentee could achieve the same

non-obviousness are satisfied, scope limits are required. To say the same thing in the reverse, only by excluding the core embodiment from patentability can we avoid the scope delineation problem.

C. *Categorical Versus Individualized Exclusion*

If the benefit of exclusions is that they permit avoidance of an administratively difficult and error-prone scope delineation inquiry, one might question whether exclusionary doctrines such as obviousness are up to the task. After all, obviousness is administratively difficult and error prone.¹⁰⁰ The answer is to divide exclusions between the categorical and the individualized. Categorical exclusions offer clear administrative cost benefits. Individualized exclusions also offer administrative cost benefits, but to a lesser extent, since they straddle the divide between rules and standards.

1. Categorical exclusions.

An exclusion is categorical when it denies patentability based on the *general nature* of the invention or the patentee, regardless of the individual merits of either. By “invention” here, I am referring to the patentee’s embodiment. Thus, if a patentee’s embodiment falls into the wrong kind of category—if it is a natural product or phenomenon—no amount of claim refinement will save it.

Almost all modern categorical exclusions are considered part of subject-matter doctrine under section 101. Conceptually, however, the two are not the same thing. It is possible to have a categorical exclusion that is not subject-matter related. One example is the categorical exclusion of inventions by non-U.S. citizens under the Patent Act of 1793.¹⁰¹ This exclusion was categorical, but it did not relate to the subject-matter sought to be patented. On the flip side of the coin, section 101 does not consist solely of categorical exclusions, as the discussion of the “abstract ideas” doctrine has shown.

In one sense, it might be difficult to distinguish between categorical exclusions and more individualized exclusions. After all, all *old*

substantive scope by using a style-conforming claim. I use the omnibus claim style here because it illustrates the substantive problem more clearly.

¹⁰⁰ See *Harries v. Air King Prods. Co.*, 183 F.2d 158, 162 (2d Cir. 1950) (Hand, J.) (describing non-obviousness as “as fugitive, impalpable, wayward, and vague a phantom as exists in the whole paraphernalia of legal concepts.”).

¹⁰¹ Patent Act of 1793, Ch. 11, § 1, 1 Stat. 318, 318 (limiting patent grants to “any person or persons, being a citizen or citizens of the United States”). This was repealed by the Patent Act of 1836, Ch. 357, § 9, 5 Stat. 117, which permitted patent applications by foreigners, but still charged them a higher application fee.

products are in some sense “categorically” excluded from patentability, too.¹⁰² The difference lies in the degree of generality of the category, and the amount of scrutiny required to establish whether an invention is categorically excluded. It is generally apparent on the face of a patent application whether the embodiment is a natural product or the inventor is a foreigner—only a minimal factual search or inquiry needs to be made. On the other hand, examining whether an embodiment is old or obvious is never self-evident on the face of the patent; at a minimum, the PTO examiner must find the relevant prior art, to prove that the patentee’s embodiment was already previously in the public domain.¹⁰³ Thus, categorical exclusions operate “at the threshold,”¹⁰⁴ barring a patent without requiring any serious examination of the invention or its individualized merits. A cursory examination—to read the patent application and determine the nature of the core embodiment described—is usually all that is required.

Because they operate at the threshold, with little individualized scrutiny, categorical exclusions are blunt instruments. This is both a good and bad thing. The good is that categorical rejections are easy to administer, since they don’t require gathering the information about prior art, nor about the economic benefits and costs of granting incremental additions to scope.¹⁰⁵ The bad is that categorical exclusions are by their nature over-inclusive. Patent applicants that create net social benefits will be denied patents simply because their work falls into the wrong category.¹⁰⁶ Such over-inclusiveness is avoided by the more narrowly-tailored inquiries of novelty, usefulness, obviousness, and proportionate scope; because no one contends that we should provide incentives for creating old, useless or obvious inventions.

¹⁰² 35 U.S.C. § 102 (2006).

¹⁰³ *In re Wilder*, 429 F.2d 447, 450 (C.C.P.A. 1970) (PTO bears burden of proving lack of novelty).

¹⁰⁴ *Diamond v. Diehr*, 450 U.S. 175, 188 (1981).

¹⁰⁵ Of which the degree of advance over the prior art is an important component. See *Eibel Process Co. v. Minn. & Ontario Paper Co.*, 261 U.S. 45, 63 (1923) (degree of scope granted by a court depends on “the real merit of the alleged discovery or invention is and whether it has advanced the art substantially”). This, of course, requires searching the prior art.

¹⁰⁶ See *In re Bergy*, 596 F.2d 952, 975 (C.C.P.A. 1979), *aff’d sub nom. Diamond v. Chakrabarty*, 47 U.S. 303 (1980) (“when [something] is new and unobvious . . . , we see no reason to deprive it . . . of the protection and advantages of the patent system by arbitrarily excluding it . . . on the sole ground that it is alive”).

2. Individual exclusions.

An individualized exclusion is one that rejects a patentee's embodiment—or rejects a particular patentee—due to their individual merits.¹⁰⁷ This is primarily reflected in patent law's criteria of novelty, usefulness, and non-obviousness. The individualized inquiry by its nature requires gathering of facts by an administrative body such as the PTO or a court, which must make a detailed search of prior art, and perhaps even test the invention to verify that it works and has utility.¹⁰⁸ An individualized inquiry also provides a measure of congruence and proportionality between the merits of an invention and whether it passes the legal test: monopolies on pre-existing or useless devices are almost by definition excessive.¹⁰⁹

Because individualized exclusions require factual scrutiny and are calibrated to the individual merits of an invention, they appear similar to scope limitations.¹¹⁰ Courts and commentators often lump the two categories together as merit-based conditions of patentability—in contradistinction to supposedly arbitrary categorical exclusions.¹¹¹ This usually makes sense, in that categorical exclusion has lower administrative cost and higher error cost than the individualized analysis that occurs during all the other stages of patent examination.

¹⁰⁷ An example of an individualized exclusion based on the merits of the patentee and not of the invention would be 35 U.S.C. § 102(f), which prohibits non-inventors (*i.e.* fraudsters and thieves) from obtaining a patent on an otherwise patentable invention. Whether a person is the true inventor, of course, requires an individualized factual inquiry.

¹⁰⁸ See *In re Kumar*, 418 F.3d 1361, 1366 (Fed. Cir. 2005) (PTO bears burden of proving obviousness); *In re Brana*, 51 F.3d 1560, 1566 (Fed. Cir. 1995) (PTO bears burden of challenging utility). Since the PTO does not have the ability to test inventions, it instead asks patent applicants to provide the results of any tests and other evidence demonstrating utility.

¹⁰⁹ See *Graham v. John Deere Co.*, 383 U.S. 1, 6 (1966) (opining that such patents would fail to promote progress).

¹¹⁰ Another reason mentioned above is that individualized exclusions are often invoked to limit scope as a practical matter. See *supra* text accompanying notes 97–99.

¹¹¹ Risch, *supra* at note 1, at 598 (arguing that historical subject-matter cases should be decided on alternative grounds, including both scope limitation and individualized exclusion doctrines); Osenga, *supra* at note 1, at 1107 (lumping novelty, non-obviousness, and scope together as “patentability requirements”); *In re Bilski*, 545 F.3d 943, 950 & n.1 (Fed. Cir. 2008) (*en banc*) (stating that section 101 is a “threshold inquiry” in contradistinction to “other legal requirements of patentability”).

At the same time, individualized exclusion only collapses into the pure-standard regime of scope limitation if we adopt the most free-floating definitions of obviousness and usefulness possible. In the case of obviousness, this is the “but for the inducement of the patent” standard; and in the case of usefulness, it is whether the invention increases net social utility in economic terms. Neither remotely describes how the legal doctrines of obviousness and utility operate in practice. Instead, we have legally “new, useful, and non-obvious” inventions that are neither socially beneficial nor likely to require any patent incentive to induce, such as methods of reducing one’s taxes—where the pre-existing incentive is self-evident and the social utility highly doubtful.¹¹² This is because non-obviousness and utility, under current law, focus primarily on the technical accomplishment to the exclusion of other considerations.¹¹³ Such doctrinal constraints in utility and non-obviousness doctrine provide structure to increase ease of application and reduce administrative cost; but they also increase rigidity and error cost. And while one could argue that obviousness and utility should become more flexible and precisely tailored to the underlying economic goal,¹¹⁴ the increased administrative cost must be accounted for, and a practical method of implementing such a heightened standard must be devised. There is little evidence that an already over-worked PTO could accommodate such increased administrative demands.¹¹⁵

¹¹² See Andrew A. Schwartz, *The Patent Office Meets the Poison Pill: Why Legal Methods Cannot be Patented*, 20 Harv. J.L. & Tech. 333, 346 (2007) (describing various patents on legal methods, including tax strategies).

¹¹³ See Merges & Duffy, *supra* at note 53, at 612 (“nonobviousness tries to measure technical, not economic, triviality”); *Juicy Whip, Inc. v. Orange Bang, Inc.*, 185 F.3d 1364, 1368 (Fed. Cir. 1999) (“The requirement of ‘utility’ in patent law is not a directive to the Patent and Trademark Office or the courts to serve as arbiters of deceptive trade practices.”).

¹¹⁴ Duffy, *supra* at note 24, at at 623; *cf.* Michael Risch, *Reinventing Usefulness*, 2010 B.Y.U. L. Rev. (forthcoming) (arguing for a “commercial utility” requirement that is similar to, but more stringent than, a requirement of economic utility).

¹¹⁵ See Chiang, *supra* at note 28, at 75-86 (proposing to reduce administrative cost by primarily relying on *ex post* litigation); Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 Nw. U. L. Rev. 1495 (2001) (arguing that the PTO should conduct only a cursory review, saving the close scrutiny for valuable patents that are litigated).

*D. Subject-Matter Doctrine Contains Both Doctrines of Scope
Limitation and Doctrines of Categorical Exclusion*

Once the policy levers of patent law are separated into the three categories of scope limitation, individualized exclusion, and categorical exclusion, the debate concerning subject-matter limitations and the excessive monopoly theory can be clarified. Critics of section 101 subject-matter limitations often point out that the blunt instrument of categorical exclusion sits uncomfortably with the goal of preventing excessive monopoly in individual patents. And categorical exclusions are indeed over-inclusive (though, as I will discuss in Part III, this does not make categorical exclusion always unjustified). But it is a mistake to collapse “section 101 doctrine” with the conceptual category of “categorical exclusions.”

The key realization from the above analysis is that although many subject-matter doctrines are categorical exclusions, section 101 subject-matter doctrine also contains what are really scope limitations, most importantly the prohibition on patenting abstract ideas and scientific principles. When discussing the theoretical policy justifications for section 101, it is important to separate out these components. This is extremely important, because almost all of the doctrinal difficulty with section 101 has concerned the abstract ideas and scientific principles doctrine.¹¹⁶ Critics of section 101 have asserted that it is both over-inclusive *and* difficult to administer: over-inclusive because it is categorical, and difficult to administer because nobody can figure out what is an abstract idea.¹¹⁷ This critique, however, conflates what in reality are two entirely distinct targets.

The Supreme Court has long held there are three categories of non-patentable subject-matter: “laws of nature, natural phenomena, and abstract ideas.”¹¹⁸ Additional exclusions have been imposed by statute

¹¹⁶ Duffy, *supra* at note 24, at 639 (“In fact, most patentable subject matter decisions are accounted for by two legal doctrines: (i) the prohibition against natural principles and natural phenomena, and (ii) the doctrine forbidding patents on abstract ideas.”). In my view, these two doctrines are one and the same in purpose and effect, and are simply limits on the scope of patents.

¹¹⁷ See *supra* notes 1–2.

¹¹⁸ *Diamond v. Diehr*, 450 U.S. 175, 185 (1981).

and administrative practice.¹¹⁹ A variety of further restrictions have existed historically, but have been overruled.¹²⁰

A close examination of all of the patentable subject-matter categories, however, reveals a clear distinction. Virtually all of the restrictions, save two, have been in the nature of categorical rules. The two exceptions are the closely-related (indeed, for all practical purposes, identical) prohibitions on patenting “laws of nature” and “abstract ideas.” The laws of nature and abstract ideas prohibitions account for almost the entirety of patentable subject-matter case law, and the consequent uncertainty in the area.¹²¹ Section 101 thus contains a dichotomy: clear exclusionary rules that have almost never caused administrative difficulty or litigation; and a vague scope limiting standard that has caused tremendous difficulty and much litigation.

1. The clarity and over-inclusiveness of subject-matter rules.

There are many categories of prohibited subject-matter that have existed at one point or another in United States patent law. Examples currently existing prohibitions include the judicial prohibition on patenting natural products and phenomena, and the statutory bans on patenting human beings and nuclear weapons. Previously accepted but overruled categories include business methods and living things.¹²²

One remarkable fact is that these exclusions were almost never subject to any significant ambiguity, despite the conventional wisdom that patentable subject-matter doctrine is “confused and inconsistent.”¹²³ Everyone has a reasonably good idea of what constitutes a “product of nature”—an apple counts, while a manufactured article doesn’t. Similarly, a “phenomenon of nature” is easy to discern. Indeed, the only Supreme Court case that concerned this doctrine is *Funk Brothers Seed Co. v. Kalo Inoculant Co.*,¹²⁴ which held, without much controversy, that bacteria in their natural state

¹¹⁹ 42 U.S.C. 2181 (2006) (nuclear weapons); Consolidated Appropriations Act of 2004, Pub. L. No. 108-199, § 634, 118 Stat. 3, 101 (2004) (cloned humans); see Merges & Duffy, *supra* at note 53, at 129-30 (discussing the PTO practice of denying patents on human-animal chimeras).

¹²⁰ *Diamond v. Chakrabarty*, 447 U.S. 303 (1980) (rejecting “living things” exclusion); *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*, 149 F.3d 1368, 1375 (Fed. Cir. 1998) (overruling the “business methods” exclusion).

¹²¹ Duffy, *supra* at note 24, at 639.

¹²² See *supra* notes 118–120.

¹²³ Risch, *supra* at note 1, at 591.

¹²⁴ 333 U.S. 127 (1948).

were products of nature and thus not patentable as such.¹²⁵ The other exclusions, such as that of human beings and atomic weapons, have engendered even less controversy.¹²⁶

The only area of any appreciable uncertainty regarding these categorical exclusions is the degree of human intervention that is required to transform a natural product into a man-made one. Ever since Learned Hand's famous 1911 decision in *Parke-Davis & Co. v. H.K. Mulford & Co.*,¹²⁷ this question has been settled by a rule that any disturbance of the natural state—even one as minimal as purification or isolation of the natural product—sufficed to transform a natural product into a man-made one.¹²⁸ While this rule is very permissive, it is not particularly vague or difficult to apply. The uncertainty in this area has resurfaced only because recent decisions have questioned the wisdom of the rule itself.¹²⁹ Legal uncertainty arising from the potential that a rule might be changed, of course, is not the same as saying the rule itself is uncertain.

The problem with these categorical exclusions has not been that their boundaries have been vague. Rather, the problem has been that they occasionally proved to be over-inclusive: that the categorical prohibition worked to exclude some socially beneficial inventions. Judge Hand's very narrow construction of the "product of nature" category in *Parke-Davis* can be understood as a reaction to this

¹²⁵ *Id.* at 130-31. *Funk Brothers* more controversially held that the combination of natural bacteria based on their natural properties of mutual non-inhibition was not patentable because that combination was obvious. *Id.* at 132. This is more controversial because all objects in the universe are ultimately combinations of natural products according to their natural properties. But that does not implicate whether the bacteria are "products of nature," which they clearly were.

¹²⁶ The prohibition on patenting human beings has never been litigated. The prohibition on patenting atomic weapons, codified at 42 U.S.C. 2181, has been addressed in only two reported decisions. *In re Brueckner*, 623 F.2d 184 (C.C.P.A. 1980); *Consol. Eng'g Corp. v. United States*, 127 F.Supp. 558 (Ct. Cl. 1955).

¹²⁷ 189 F. 95 (S.D.N.Y. 1911), *aff'd*, 196 F. 496 (2d Cir. 1912).

¹²⁸ *Id.* at 103; see also U.S. Patent and Trademark Office, *Utility Examination Guidelines*, 66 Fed. Reg. 1092, 1093 (Jan. 5, 2001) ("compounds isolated from nature are patentable").

¹²⁹ *Intervet Inc. v. Merial Ltd.*, 2010 WL 3064311, at *11 (Fed. Cir. 2010) (Dyk, J., dissenting) (questioning *Parke-Davis* rule by arguing that "[i]t would be difficult to argue . . . that one could patent the leaves of a plant merely because the leaves do not occur in nature in their isolated form"); *Association for Molecular Pathology v. U.S. Patent and Trademark Office*, 2010 WL 1233416 (S.D.N.Y. 2010) (dismissing *Parke-Davis* rule as dicta).

problem, since he placed much emphasis on the social benefit of purified adrenaline in devising his rule.¹³⁰ The judicial concern with over-inclusiveness is also aptly illustrated by the demise of the “living things” exclusion upon the advent of genetic engineering.

Historically, living things were regarded as non-patentable subject-matter, as a corollary of the prohibition on patenting natural products.¹³¹ In this period, the rule was not over-inclusive: few would argue that natural products should be patented, since natural products are always pre-existing,¹³² and historically all living things were products of nature. But with the advent of genetic engineering, it became possible to create new, useful, and non-obvious organisms that were, at the same time, alive. This created a problem: if patents were fundamentally utilitarian, it seemed perverse that a patent on a new, useful, and non-obvious thing should be denied simply because it fell into the “wrong” category.

Courts quickly began to see the “living things” exception an arbitrary limit serving no useful policy goal, and eliminated the exception. Writing for the Court of Customs and Patent Appeals, Judge Giles Rich articulated the over-inclusiveness concern explicitly:

[M]icroorganisms have long been important tools in the chemical industry, especially its pharmaceutical branch, and when such a useful, industrial tool is invented which is new and unobvious, so that it complies with those conditions for patentability, we see no reason to deprive it or its creator or owner of the protection and advantages of the patent system by arbitrarily excluding it at the outset from the § 101 categories of patentable invention on the sole ground that it is alive.¹³³

When the issue reached the Supreme Court, that Court largely agreed, eliminating the living things exception and relying instead on the novelty and utility requirements to police patentability: “the patentee has produced a new bacterium with markedly different characteristics from any found in nature, and one having the potential

¹³⁰ *Parke-Davis*, 189 F. at 114-15.

¹³¹ Duffy, *supra* at note 24, at 629-30.

¹³² See *Merck & Co. v. Olin Mathieson Chem. Corp.*, 253 F.2d 156 (4th Cir. 1958) (arguing that “product of nature” rejections are really based on obviousness or lack of novelty). It should be noted that a product of nature can be technically “novel” in the eyes of patent law, because patent law defines novelty as not having previously been in public use *within the United States*. See 35 U.S.C. § 102(a) & (b) (2006).

¹³³ *In re Bergy*, 596 F.2d 952, 974 (C.C.P.A. 1974).

for significant utility. His discovery is not nature's handiwork, but his own; accordingly it is patentable subject matter under § 101.”¹³⁴

The sum of the above is that categorical exclusions under section 101 have not caused significant vagueness or administrability concerns, but they have sometimes caused over-inclusiveness concerns. Those over-inclusiveness concerns, in turn, have led to the demise of many such exclusions from current law.

The effect of over-inclusiveness concerns is on vivid display in the elimination of the most significant—and the most controversial—categorical exclusion, that of methods of doing business. For most of the history of the patent system, this exclusion was well-accepted.¹³⁵ Despite the potential fuzziness of what might constitute a business method, the vagueness of the category does not appear to have posed significant administrative problems, or at least any boundary-defining problems do not appear in reported cases.¹³⁶ Rather, the problem was that the boundaries were clear enough to show that many worthwhile business-related inventions fell within the exclusion and were being eliminated from patentability. As Giles Rich observed in 1960:

Invaluable though it may be to individuals, the public, and national defense, the invention of a more effective organization of the materials in, and the techniques of teaching a course in physics, chemistry, or Russian is not a patentable invention Also outside that group is one of the greatest inventions of our times, the diaper service.¹³⁷

¹³⁴ *Diamond v. Chakrabarty*, 447 U.S. 303, 310 (1980).

¹³⁵ See John R. Thomas, *The Patenting of the Liberal Professions*, 40 B.C. L. Rev. 1139, 1145-47 (1999) (describing history).

¹³⁶ See *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*, 149 F.3d 1368, 1375-76 (Fed. Cir. 1998) (arguing that nearly all the reported cases involving the business method exclusion in fact turned on other patentability criteria).

¹³⁷ Giles S. Rich, *Principles of Patentability*, 28 Geo. Wash. L. Rev. 393, 393-94 (1960); see also Giles S. Rich, *The Relation Between Patent Practices and the Anti-Monopoly Laws—Part II*, 24 J. Pat. Office Soc’y 159, 160-61 (1942) (“By far the greater portion of human inventions and discoveries lie outside the domain prescribed by the patent statutes, which limit the grant of patents, generally speaking, to the industrial arts. For example no protection is afforded to the novel, useful, and clearly inventive schemes of conducting business . . . or to pedagogical and other sociological improvements of the highest value.”)

Nearly forty years later, Judge Rich personally corrected this perceived defect of the patent system, by overturning the business method exclusion in the famous case of *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*¹³⁸

2. The judicial struggle to define abstract ideas.

In contrast to relatively clear and bright-line categories such as natural products and phenomena, the courts have had tremendous difficulty with the other line of subject-matter doctrine: the prohibition on patents covering “laws of nature” and “abstract ideas.” In many ways, this is because courts have confused the subject-matter question of what kinds of patentee creations can be patented, versus the scope delineation question of how much broader than precise replicas of a patentee’s creation the patent monopoly should reach. Instead of *patentability*—i.e. whether a patent can issue at all—the “laws of nature” and “abstract ideas” doctrines are fundamentally matters of *scope*.

The prohibition on patenting abstract ideas and natural principles is neither an exclusion nor a rule. It is not an *exclusion* from patentability because no patentee will ever seek to enter the patent system with just an abstract idea in the first place. As John Duffy notes: “Quite apart from § 101 of the Patent Act, an abstract idea could not be patented because abstraction is the very antithesis of the precision required by the disclosure provisions of the Patent Act.”¹³⁹ A patentee describing *only* an abstract idea in the specification could not describe how to “make and use” his invention;¹⁴⁰ nor show how the abstract idea is “useful” in the practical sense that patent law requires.¹⁴¹ As a doctrine of exclusion, it would be completely redundant, and it has never been so understood or applied.

Instead, the effect of the abstract idea doctrine is to delineate proper scope. This is demonstrated by the very first case that applied the doctrine, *O’Reilly v. Morse*,¹⁴² concerning Samuel Morse’s telegraph. The precise embodiment in the specification was an early telegraph machine using electric current.¹⁴³ From this embodiment, Samuel

¹³⁸ 149 F.3d at 1375 (“We take this opportunity to lay this ill-conceived exception to rest.”).

¹³⁹ Duffy, *supra* at note 24, at 645.

¹⁴⁰ 35 U.S.C. § 112 (2006).

¹⁴¹ See *Brenner v. Mason*, 383 U.S. 519, 534-35 (1966) (defining utility to require “specific benefit . . . in currently available form”).

¹⁴² 56 U.S. (15 How.) 62 (1853).

¹⁴³ See U.S. Patent No. 1,647 (issued June 20, 1840).

Morse filed eight claims of varying breadth, reflecting the different levels of abstraction at which a telegraph can be described. One rather narrow claim hewed closely to the precise contours of a telegraph machine:

Third. I [] claim as my invention and improvement the combination of machinery herein described, consisting of the generation of electricity, the circuit of conductors, the contrivance for closing and breaking the circuit, the electro-magnet, the pen or contrivance for marking, and the machinery for sustaining and moving the paper, altogether constituting one apparatus of telegraphic machinery, which I denominate the American Electro-Magnetic Telegraph.¹⁴⁴

Because a narrow claim to electro-magnetic telegraph machines might be evaded, however, Morse also sought a far more abstract and broad claim:

Eighth. I do not propose to limit myself to the specific machinery or parts of machinery described in the foregoing specifications and claims, the essence of my invention being the use of the motive power of the electric or galvanic current, which I call electro-magnetism, however developed, for making or printing intelligible characters, letters, or signs, at any distances.¹⁴⁵

The Supreme Court easily found that a telegraph machine was new, useful, and non-obvious, and sustained Morse's first seven claims.¹⁴⁶ The Court, however, invalidated the eighth claim, on the theory that the "motive power" of electro-magnetism was not itself patentable.¹⁴⁷ Although *Morse* did not itself use the phrase "law of nature" to describe the principles of electromagnetism, the decision has been continuously cited for the proposition that laws of nature and abstract ideas are not patentable.¹⁴⁸

¹⁴⁴ *Morse*, 56 U.S. at 85-86.

¹⁴⁵ *Id.* at 86.

¹⁴⁶ *Id.* at 112.

¹⁴⁷ *Id.* at 114 (drawing the analogy to printing presses and stating that "it has never, we believe, been supposed by anyone that the first inventor of a steam printing press was entitled to the exclusive use of steam as a motive power, however developed, for marking or printing intelligible characters").

¹⁴⁸ *Lab. Corp. of Am. Holdings v. Metabolite Labs., Inc.*, 548 U.S. 124, 135 (2006) (Breyer, J., dissenting from dismissal of writ of certiorari); *In re*

It is clear, however, that *Morse* applied the abstract idea doctrine only to limit the *scope* of the telegraph patent; not to exclude anything from patentability. The most telling sign is that Samuel Morse received a valid patent on his telegraph embodiment and successfully enforced it.¹⁴⁹ Thus, the doctrine did not have the effect of depriving him a patent. Moreover, the policy concerns expressed by the Court were not directed to patentability, but only to scope. In particular, the Court was concerned that a broad claim would impede future variants that were very different from Morse's specific embodiment:

For aught that we now know, some future inventor, in the onward march of science, may discover a mode of writing or printing at a distance by means of the electric or galvanic current, without using any part of the process or combination set forth in the plaintiff's specification. His invention may be less complicated -- less liable to get out of order -- less expensive in construction, and in its operation. But yet if it is covered by this patent, the inventor could not use it, nor the public have the benefit of it, without the permission of this patentee. . . . In fine, [Morse] claims an exclusive right to use a manner and process which he has not described and indeed had not invented, and therefore could not describe when he obtained his patent. The court is of opinion that the claim is too broad, and not warranted by law.¹⁵⁰

The result of this focus on scope, not surprisingly, is that narrower claiming cured the problem; and the narrower claims were all allowed. In this vein, it is particularly noteworthy that the narrower claims still covered "ideas" in every sense of the word. For example, Morse's Fifth claim, upheld by the Court, stated:

I claim as my invention the system of signs, consisting of dots and spaces, and of dots, spaces, and horizontal lines, for numerals, letters, words, or sentences, substantially as herein set forth and illustrated, for telegraphic purposes.¹⁵¹

This claim covered, in essence, Morse code. It was abstract and required no specific machinery. The contrast between the Fifth and

Chatfield, 545 F.2d 152, 157 (C.C.P.A. 1976); see also *Gottschalk v. Benson*, 409 U.S. 63, 67-68 (1972).

¹⁴⁹ *Morse*, 56 U.S. at 124 (finding infringement).

¹⁵⁰ *Id.* at 113.

¹⁵¹ *Id.* at 86.

Eighth claim illustrates that the problem is not that the Eighth claim covered an idea or that it was an abstraction, but only that it was *excessively* abstract.

The second point to appreciate about the abstract idea doctrine is that “excessiveness” is highly fact-specific and malleable. This becomes clear when *Morse* is contrasted with the next important Supreme Court case on the subject, concerning Alexander Graham Bell’s patent on the telephone.¹⁵² Bell’s patent contained a claim that was, for practical purposes, almost identical to Morse’s eighth claim:

5. The method of and apparatus for transmitting vocal or other sounds telegraphically, as herein described, by causing electrical undulations similar in form to the vibrations of the air accompanying the said vocal or other sounds, substantially as set forth.¹⁵³

In plain English, this claim had the legal effect of covering all use of electrical current to transmit sound, where the current went up and down in patterns corresponding to sound vibrations.¹⁵⁴ Like *Morse*, it was not limited to any particular telephone machine, and covered later-developed telephones that were “less complicated -- less liable to get out of order -- less expensive in construction, and in [their] operation.” In contrast to *Morse*, however, the Supreme Court upheld this claim.¹⁵⁵ The juxtaposition of these two cases makes clear that determining when something is an “abstract idea” is not easily reducible to a bright-line rule.¹⁵⁶

The analysis is identical when we adopt the older formulation of prohibiting claims covering “laws of nature” or a “natural principle.” The label of a “law of nature” is, like the label of an “abstract” idea, an infinitely malleable and thus conclusory label. Applied literally, *everything* in the universe is the result of manipulating matter

¹⁵² *The Telephone Cases*, 126 U.S. 1 (1888).

¹⁵³ *Id.* at 13-14.

¹⁵⁴ In patent law, the terms “as herein described” and “substantially as set forth” are regarded as meaningless fluff when used in a claim. *Dugan v. Ford Instrument Co.*, 15 F. Supp. 442, 445 (E.D.N.Y. 1936). Thus, Bell’s claim encompassed later-developed variants that differed significantly from the specific telephone disclosed in his patent specification.

¹⁵⁵ *Telephone Cases*, 126 U.S. at 533 (“The patent for the art does not necessarily involve a patent for the particular means employed for using it.”).

¹⁵⁶ See George Ticknor Curtis, *A Treatise on the Law of Patents* § 166 (2d ed. 1873); see also Duffy, *supra* at note 24, at 642-44.

according to the laws of nature such as the laws of physics,¹⁵⁷ just as every patent covers an abstract idea in the colloquial sense of the word. The claimed idea behind a patented embodiment is always abstract; and it is always reducible to some combination of natural principles.¹⁵⁸ The vacuousness of the standard was recognized by Justice Frankfurter in *Funk Brothers Seed Co. v. Kalo Inoculant Co.*:¹⁵⁹

[S]uch terms as "the work of nature" and the "laws of nature . . . are vague and malleable terms infected with too much ambiguity and equivocation. Everything that happens may be deemed "the work of nature," and any patentable composite exemplifies in its properties "the laws of nature." Arguments drawn from such terms for ascertaining patentability could fairly be employed to challenge almost every patent.¹⁶⁰

Thus, the prohibition against patenting "laws of nature" cannot be applied literally. Rather, like the prohibition on patenting abstract ideas, a claimed idea becomes a prohibited "law of nature" only when it is somehow "excessively" broad.

Finally, it is important to note that excessiveness in this context is determined by a pragmatic balancing of real-world consequences. This is aptly demonstrated by the Supreme Court's recent decision in *Bilski v. Kappos*.¹⁶¹ The unanimous Court (on this point) concluded that Bilski's claims to hedging risk in commodities transactions were an abstract idea. In articulating precisely *why*, however, the Court focused not on semantic content but real-world commerce:

Claims 1 and 4 in petitioners' application explain the basic concept of hedging, or protecting against risk: Hedging is a fundamental economic practice long prevalent in our system of commerce and taught in any introductory finance class. . . . Allowing petitioners to patent risk hedging would pre-empt use of this approach

¹⁵⁷ See Immanuel Kant, *Critique of Pure Reason* 302 (J.M.D. Meiklejohn trans., Prometheus Books 1990) (1787).

¹⁵⁸ 1 William C. Robinson, *The Law of Patents for Useful Inventions* § 134 (1890) ("No proposition has been more frequently or positively stated by the courts than that a principle is not a patentable invention, and yet with almost equal positiveness and frequency they have declared that the subject matter covered by a patent is the principle of the invention.").

¹⁵⁹ 333 U.S. 127 (1948).

¹⁶⁰ *Id.* at 134-35 (Frankfurter, J., concurring).

¹⁶¹ 130 S. Ct. 3218 (2010).

in all fields, and would effectively grant a monopoly over an abstract idea.¹⁶²

Aside from the *ipse dixit* statement that allowing the claim would “grant a monopoly over an abstract idea,” the basis of the Court’s decision seems to be that hedging is (1) fundamental, and (2) long-prevalent. The fact that hedging is “fundamental” means that a patent is likely to generate high monopoly cost, and the Court’s implicit finding that it is pre-existing negates all of Bilski’s contribution, rendering the cost-benefit balance decisively unfavorable and thus the monopoly cost excessive.

A similar pragmatic balancing can be seen in the older cases. The best explanation for the contrast between *Morse* and the *Telephone Cases* is not that transmitting sounds via electric current is somehow more or less “abstract” than transmitting characters, but that Bell’s practical need for broad patent protection was greater than Morse’s.¹⁶³ *Gottschalk v. Benson*,¹⁶⁴ which held that a computer for converting binary-coded decimal numbers to pure binary numbers was an abstract idea,¹⁶⁵ likewise emphasized the broad monopoly that would result from covering “both known and unknown uses of the BCD to pure binary conversion . . . [,] vary[ing] from the operation of a train to verification of drivers’ licenses to researching the law books for precedents.”¹⁶⁶

The sum of the above is that the abstract ideas doctrine is a scope limiting standard, which is difficult for judges to administer. The cost-benefit balance is different for every invention, so prior cases are not very useful for determining whether scope is excessive, and in any event the *United States Reports* and the *Federal Reporter* do not often give a detailed record the social cost-benefit balance of an invention, so

¹⁶² Id. at 3231 (internal quotations and citations omitted).

¹⁶³ See 1 Donald S. Chisum, *Chisum on Patents* § 1.03[2][d] (2009) (“The holding of the Telephone Case was undoubtedly influenced by the desire to secure recognition and reward for Alexander Graham Bell. Striking the [Bell] claim would have emasculated the basic telephone patent. Striking Morse’s eighth claim, on the other hand, did not prevent Morse from obtaining effective relief.”).

¹⁶⁴ 409 U.S. 63 (1972).

¹⁶⁵ A “binary-coded decimal” is a decimal number where each digit is represented by a four digit binary number. For example, the decimal number “47” would be represented in binary digit form as “0100 0111”; since “0100” is the binary form of “4” and “0111” is the binary form of “7.” The pure binary equivalent of 47 would be 101111.

¹⁶⁶ *Benson*, 409 U.S. at 68. It should be noted that *every* patent covers “known and unknown uses.” Feldman, *supra* at note 61, at ¶ 60.

the tools of the judicial craft are ill-adapted to the task. At the same time, scope limiting standards are still necessary in patent law because they permit precise tailoring, and these benefits and drawbacks of the abstract ideas doctrine are shared with other scope limiting doctrines such as enablement, which judges still implement despite all the problems. The point is not to say that the abstract ideas doctrine is superior or inferior to section 101 categorical exclusions such as the natural products doctrine, but simply to note that they are fundamentally different creates occupying diametrically opposite positions on the rules-standards continuum. There is no unified “section 101” subject-matter doctrine or test.

E. The Importance of Distinguishing Categorical Exclusion from Scope Limitation

1. Confused doctrine: *In re Bilski*.

The doctrinal problems that can arise by treating section 101 as a unified whole is aptly demonstrated by the Federal Circuit’s recent *en banc* decision in *In re Bilski*.¹⁶⁷ After reviewing the history of section 101, *Bilski* attempted to lay down a unitary test for patentable subject-matter: a claim is *per se* patentable subject-matter if it is “tied to a particular machine, or . . . transforms an article,”¹⁶⁸ and not otherwise.¹⁶⁹

It should be immediately evident that this “machine-or-transformation” rule is primarily concerned with the *scope*, not *patentability*. aspects of section 101. This is because the machines and transformations are utterly redundant as far as inventions-as-embodiments are concerned. All patented and infringing embodiments necessarily have tangible effects in the real world, and otherwise it would be impossible to establish utility or to detect infringement. Thus, inventions that have no physical transformations at all are ones that no one would care about. The only question is whether the transformation is *included in a claim*, or is instead abstracted out.

By requiring that claim recite a machine-or-transformation, *Bilski*’s test is misguided and unworkable for two reasons. For section 101 categorical exclusions, the machine-or-transformation test is misguided because the test can only narrow scope and never denies patentability (since all embodiments ultimately effect physical transformations), while categorical exclusion doctrines seek to preclude

¹⁶⁷ 545 F.3d 943 (Fed. Cir. 2008) (en banc).

¹⁶⁸ *Id.* at 961.

¹⁶⁹ *Id.* at 959 (emphasizing the “machine-or-transformation” test is the sole test).

patentability altogether in order to save administrative cost. For section 101 scope limitations, the test is misguided because it imposes a rigid rule, and rigid rules about scope have never escaped the problem of unacceptable over- and under-inclusiveness.

The first harm is easily demonstrated by the hypothetical example of Einstein applying for a patent on nuclear fusion, the precise embodiment being the physical process occurring in the Sun. This is, without any doubt, an unpatentable “natural phenomenon” under traditional doctrine.¹⁷⁰ At the same time, the process of nuclear fusion clearly “transforms an article,” since it converts hydrogen to helium. Nor can the patent be rejected on simple novelty grounds, since although nuclear fusion was previously in public use, the use was not within the United States as required by the novelty statute.¹⁷¹ In short, the implication of *Bilski* is that Einstein should have been able to obtain a patent on the process of nuclear fusion; and that future scientists would be able to obtain patents on natural processes such as black holes.¹⁷² The absurdity of applying the “machine-or-transformation” test to natural processes¹⁷³—which by definition involve transformations at some point—becomes apparent.

The second harm of *Bilski* lies in imposing a rigid test on the scope-limitation aspects of section 101. On some level, of course, predictable rules for patent scope would be highly desirable. The only problem is that such predictability has always come at too high a price in terms of over- and under-inclusiveness when applied in the context of patent scope, and *Bilski* gives no reason at all to think it will escape this history.¹⁷⁴ After all, predictability in patent scope would be served by

¹⁷⁰ Note that it is not a “law of nature,” since I am focusing on the physical process rather than scientific principles underlying the process.

¹⁷¹ 35 U.S.C. § 102(b) (2006).

¹⁷² One argument might be that the scientist did not “invent” the natural process, and thus would be excluded under 35 U.S.C. § 102(f), which excludes a patent if the applicant “did not himself invent the subject matter sought to be patented.” This would require a radical reinterpretation of section 102(f), which currently is understood as a narrow anti-theft provision. *New England Braiding Co. v. A.W. Chesterton Co.*, 970 F.2d 878, 883 (Fed. Cir. 1992). Indeed, such a broad understanding of section 102(f) would require it to swallow up the rest of section 102 and 103, in that a patentee would not be the true “inventor” of something that was old or obvious.

¹⁷³ See *Bilski*, 545 F.3d at 952 n.5 (explicitly including “natural phenomena” within the machine-or-transformation test).

¹⁷⁴ *Royal Typewriter Co. v. Remington Rand, Inc.*, 168 F.2d 691, 693-94 (2d Cir. 1948) (Hand, J.) (“courts have differed, and always will differ, as to the allowable latitude in a given instance”).

consistently enforcing a rule that said patent scope shall be limited to slavish reproduction of the patentee's embodiment; or equally by enforcing the opposite rule that patent scope shall be unlimited. But both produce equally unacceptable outcomes, which is why neither of these rules have ever been seriously enforced, though both have been articulated at one point or another.¹⁷⁵

The machine-or-transformation rule is both over- and under-inclusive. The rule is very permissive (i.e. under-inclusive) in the sense that it permits any claim that recites a physical transformation, even to claims over natural phenomena.¹⁷⁶ At the same time, the rule is over-inclusive because it limits the scope of all patents to a particular physical implementation, which for some pioneering patents will be inefficiently narrow. These over- and under-inclusiveness problems, however, are at present largely hidden by the vagueness of the test. The vagueness lies in what constitutes tying a claim to a "particular" machine or transformation. To the extent this can be met very loosely—such as by tying a claim to a generic "computer"—the machine-or-transformation test will impose no meaningful limit on patent scope. To the extent that the particularity requirement mandates specifying a particular machine down to its last bolt, nut, and paint color, and a particular transformation down to the last atom, then the test will reduce patent scope to covering only slavish copying, which minimizes the under-inclusiveness problem but maximizes the over-inclusiveness problem by rendering patents practically worthless. And to the extent courts simply oscillate between demanding high and low degrees of particularity in patent claims, the test will suffer less from these over- and under-inclusiveness problems, but lose all the predictability that was the point of creating the machine-or-transformation rule in the first place.

The Supreme Court's recent opinion affirming the Federal Circuit in *Bilski v. Kappos*¹⁷⁷ mitigates some of the problems that I have identified, but leaves much of the doctrinal chaos in place. The Court held that there are three judge-made exceptions to patentable subject-matter: laws of nature, natural phenomena, and abstract ideas;¹⁷⁸ and

¹⁷⁵ See *supra* text accompanying notes 70–75.

¹⁷⁶ See *Prometheus Labs., Inc. v. Mayo Collaborative Servs.*, 581 F.3d 1336, 1349 (Fed. Cir. 2009) ("because the claims meet the machine-or-transformation test, they do not preempt a fundamental principle"), *vacated and remanded*, 2010 WL 2571881 (U.S.); see also *supra* text accompanying notes 170–173.

¹⁷⁷ 130 S. Ct. 3218 (2010).

¹⁷⁸ *Id.* at 3225.

that the machine-or-transformation test provides a “useful clue” to whether something satisfies section 101, but is not definitive in that regard.¹⁷⁹

The Court’s decision marks an improvement over the Federal Circuit in at least two respects. First, the demotion of the machine-or-transformation test from a binding rule to a “useful clue” mitigates the rigidity (*i.e.* the over- and under-inclusiveness) of the test. Second, the reaffirmation that there are *three* distinct exceptions to section 101 makes it more difficult for the Federal Circuit to devise another all-encompassing test merging the scope-limiting and category excluding aspects of section 101 in the future. Beyond these narrow improvements, however, much doctrinal confusion will remain because the Court provided little guidance on what the established exceptions mean, and also did not foreclose new exceptions in the future.¹⁸⁰

2. Confused academic discussion.

The effect of confusing the categorical exclusions of section 101 with the scope limitations of section 101 is the contradictory argument noted at the beginning of this Article. By treating section 101 as a unitary doctrine, critics assail it as both overly rigid and excessively malleable at the same time.

Once we make a distinction between categorical exclusion and scope limitation, however, section 101 makes more sense. Contrary to the critique of section 101 as generally over-inclusive,¹⁸¹ the scope limitation aspect of section 101 under the abstract ideas doctrine poses no problems of over- or under-inclusiveness. The definition of “abstract” is malleable and simply means that courts find the resulting patent scope to be too broad.¹⁸² In this way, it is largely redundant with current enablement doctrine in effect.¹⁸³ Indeed, *Morse* is often taught as an enablement case in law school patent law courses.¹⁸⁴

¹⁷⁹ *Id.* at 3227.

¹⁸⁰ *Id.* at 3231 (“In disapproving an exclusive machine-or-transformation test, we by no means foreclose the Federal Circuit's development of other limiting criteria that further the purposes of the Patent Act and are not inconsistent with its text.”).

¹⁸¹ See, e.g., Peter Lee, *Contracting to Preserve Open Science: Consideration-Based Regulation in Patent Law*, 58 Emory L.J. 889, 914 (2009) (“limitations on patentable subject matter [] would be overinclusive and undermine private incentives to invent and develop research tools”);

¹⁸² See *supra* Part II.D.2.

¹⁸³ See *supra* text accompanying notes 81–86

¹⁸⁴ See Craig Allen Nard, *The Law of Patents* 50-54 (2008) (providing *Morse* as an enablement case).

Of course, if the abstract idea doctrine is redundant with enablement, that is also an argument to fold the doctrine into enablement order to simplify patent law. Indeed, one practical benefit from such a move would be to solve an “order of battle” problem. Ideally, scope delineation should be the *last* exercise performed by a court or the PTO, because it is the most complicated and administratively expensive inquiry.¹⁸⁵ However, section 101 issues are conventionally resolved *first* under PTO and judicial practice, when the least information is available.¹⁸⁶ Thus, section 101 does not appear to be a fit doctrinal home for the scope delineation exercise.

At the same time, retaining the abstract idea doctrine as an independent scope limitation rooted in section 101 has two advantages. The first is that the doctrine is long-standing, and abolishing it would thus be disruptive. The second is that the phrasing of the prohibition as one on claiming “abstract ideas” helpfully directs the inquiry to the levels of abstraction problem I have described. In contrast, the language of “enablement” naturally suggests the inquiry as one of whether people in the field can make and use all of the infringing products being claimed.¹⁸⁷ As Jeffrey Lefstin and myself have demonstrated, this superficial “make and use” standard is useless in delineating scope.¹⁸⁸ Every embodiment in the universe falls into one of three categories: (1) pre-existing embodiments known before filing; (2) embodiments taught by the patent at filing; and (3) after-arising embodiments. Category 1 is clearly unpatentable on novelty grounds; while category 2 is extremely narrow. Thus, the question of scope is the extent to which patents may cover accused products falling into category 3. At the time of patent filing, after-developed technology by definition cannot be made and used; while by the time of patent

¹⁸⁵ See Risch, *supra* at note 1, at 651 (“If § 101 determinations are to be made first, then such determinations should be straightforward and categorical.”).

¹⁸⁶ *Bilski*, 545 F.3d at 950 n.1 (“[G]iven that § 101 is a threshold requirement, claims that are clearly drawn to unpatentable subject matter should be identified and rejected on that basis. Thus, an examiner should generally first satisfy herself that the application's claims are drawn to patent-eligible subject matter.”). It should be emphasized that no doctrine requires this order of battle. *Id.* (“if the examiner deems it appropriate, she may reject the claim on any other ground(s) without addressing § 101”).

¹⁸⁷ Merges & Nelson, *supra* at note 71, at 845 (“Under section 112, the disclosure must be sufficient to enable someone skilled in the art to make and use all the embodiments of the invention claimed in the patent.”).

¹⁸⁸ Lefstin, *supra* at note 81, at 1166 (“an inventor need not—an in most cases can not [sic]—enable all things falling within the scope of his claim”); Chiang, *supra* at note 9, at 17-20. See *supra* text accompanying notes 9–75.

infringement, an after-developed accused product almost inevitably can be made and used, because an ordinary person in the field—the infringer—has just made and used it.¹⁸⁹ And choosing some point in-between patent filing and infringement to make the determination raises the problem of arbitrariness.

Although the “whether a person in the field can make and use the infringing product” approach simply does not work, it remains fashionable to describe enablement doctrine in these terms.¹⁹⁰ It is likely that the linguistic label of “enablement” has confused the issue. The label of “abstract idea” does not carry the same baggage.¹⁹¹

The other aspect of section 101—its categorical exclusions—require a different analysis. Because of their tendency toward over- and under-inclusiveness, categorical exclusions are not very adept instruments at preventing excessive monopoly in individual cases, making the excessive monopoly theory problematic in this context. Rather, if categorical exclusions are to be justified, the excessive monopoly theory must be applied at a more systemic level, after considering the administrative cost of the patent system as a whole.

III. AN ADMINISTRATIVE COST THEORY OF CATEGORICAL EXCLUSIONS

The bluntness of categorical exclusions, with the reduced administrative cost but greater tendency toward over-inclusiveness error, reflects a rules-versus-standards debate. Categorical exclusions are analogous to bright-line rules: easy to administer, predicable *ex ante*, but prone to over- and under-inclusiveness. Individualized examination, relatively speaking, becomes analogous to more fuzzy standards: tailored and thus congruent and proportional to a policy goal, but more costly to administer and less predictable *ex ante*. Whether rules or standards are preferable, and in what circumstances,

¹⁸⁹ On rare occasions, one might argue that the infringer has extraordinary skill and thus would not reflect the abilities of an ordinary person in the field. But extraordinary skill, by definition, is rare. Requiring extraordinary flashes of genius to escape infringement would thus make almost every subsequent device infringing. *Cf. Graham v. John Deere Co.*, 383 U.S. 1, 15 n.7 (1966) (repudiating the “flash of creative genius” test for patentability).

¹⁹⁰ See, e.g., Timothy R. Holbrook, *Equivalency and Patent Law's Possession Paradox*, 23 Harv. J.L. & Tech. 1, 40 (2009) (advocating a test of whether the “accused device [is] enabled by the patent at the time of infringement”); Merges & Nelson, *supra* at note 71, at 845.

¹⁹¹ *Cf. Nash v. CBS, Inc.*, 899 F.2d 1537, 1540 (7th Cir. 1990) (“Sometimes called the ‘abstractions test’, Hand’s insight is not a ‘test’ at all. It is a clever way to pose the difficulties that require courts to avoid either extreme of the continuum of generality.”).

is the subject of an enormous literature, which will only be briefly reviewed here.¹⁹²

A. Rules Versus Standards

As has been frequently recognized, legal directives can be formulated as bright-line rules or flexible standards.¹⁹³ Both rules and standards aim to serve some underlying policy goal, the difference lies in the amount of discretion and complexity that occurs in *ex post* decision-making.¹⁹⁴ A rule attempts to capture the background policy goal when it is proclaimed, but once proclaimed it operates independently, even when its adherence produces results that are contrary to the original policy goal.¹⁹⁵ For example, a speed limit of 55 mph is a rule. It is proclaimed to promote traffic safety. But once proclaimed, a driver that exceeds 55 mph violates the law even on a empty highway where exceeding 55 mph poses no risk to traffic safety. In this way, rules are necessarily over- and under-inclusive in their attempt to capture the underlying policy goal.

In contrast to rules, a legal directive is standard-like when it maps an *ex post* decision directly onto the underlying policy goal.¹⁹⁶ Vague legal directives such as judging negligence by reference to “reasonableness” are prototypical standards. Standards reduce the problems of over- and under-inclusiveness, but create greater problems of *ex ante* notice, and require greater cost *ex post* to enforce.¹⁹⁷ A legal directive for drivers to “drive safely,” for example, would reduce the over- and under-inclusiveness of a rigid speed limit. But it would be very difficult and expensive to administer, because a judge would have to collect large amounts of information on the road condition, the surrounding geography, the speed, the amount of daylight, etc., and even with such complete information people may still disagree about

¹⁹² See generally Kaplow, *supra* at note 94; Colin S. Diver, *The Optimal Precision of Administrative Rules*, 93 Yale L.J. 65 (1983); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 Harv. L. Rev. 1685, 1687-1713 (1976); Ronald M. Dworkin, *The Model of Rules*, 35 U. Chi. L. Rev. 14, 22-29 (1967).

¹⁹³ Pierre Schlag, *Rules and Standards*, 33 UCLA L. Rev. 379, 381 (1985).

¹⁹⁴ Kathleen M. Sullivan, *Foreword: The Justices of Rules and Standards*, 106 Harv. L. Rev. 22, 57 (1992) (“These mediating legal directives take different forms that vary in the relative discretion they afford the decisionmaker.”).

¹⁹⁵ *Id.* at 58.

¹⁹⁶ *Id.*

¹⁹⁷ Kaplow, *supra* at note 94, at 577 (“Rules cost more to promulgate; standards cost more to enforce.”).

what constitutes “safe” driving in any particular circumstance, as some people prefer more cautious driving than others.

It should be noted that few legal directives are “pure rules” or “pure standards,”¹⁹⁸ and most legal directives take shades of one or the other. Pure rules are non-discretionary and factually simple: a defendant is certainly guilty if he exceeds the speed limit, and speed is the only relevant fact. Pure standards are highly discretionary and factually complex, in that “safe” or “reasonable” driving requires consideration of an almost infinite array of facts; and even if we had no factual dispute, different people would *still* disagree about what constitutes “safe” driving on the same facts.¹⁹⁹ Some legal directives, however, fall in between, or mix the two factors, and are thus difficult to classify. For example, the tax code is designed to be rigid—it seeks to permit no discretion as to the amount of tax liability on a given set of facts—but is so factually complex that, practically speaking, there is usually a great deal of vagueness about tax liability for large corporations.²⁰⁰ Similarly, the *Carroll Towing* formula of $B < PL$ can be viewed as rigid rule that allows no legal discretion: always find negligence if the *Carroll Towing* formula is satisfied.²⁰¹ But the facts required for the

¹⁹⁸ Margaret Jane Radin, *Presumptive Positivism and Trivial Cases*, 14 Harv. J.L. & Pub. Pol’y 823, 828-32 (1991) (explaining that rules and standards are endpoints on a continuum).

¹⁹⁹ See Dan Kahan, David A. Hoffman, and Donald Braman, *Whose Eyes are You Going to Believe? Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 Harv. L. Rev. 837 (2009) (showing that different people have different assessments of the safety of driving shown by the video in *Scott v. Harris*, 550 U.S. 372 (2007)).

²⁰⁰ Isaac Ehrlich & Richard Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. Legal Stud. 257, 258 (1974) (“a rule that required the weighing of many circumstances . . . would be like a standard”); see also Kyle D. Logue, *Tax Law Uncertainty and the Role of Tax Insurance*, 25 Va. Tax Rev. 339, 369-70 (2005) (“a system that consists almost entirely of complex rules also can produce a type of legal uncertainty”).

²⁰¹ *United States v. Carroll Towing Co.*, 159 F.2d 169, 173 (2d Cir. 1947) (“if the probability be called P; the injury, L; and the burden, B; liability depends upon whether B is less than L multiplied by P”); Ehrlich & Posner, *supra* at note 200, at 258 “The efficiency standard itself could be regarded as a rule of social choice designed to implement a broader standard.”); see also Richard A. Posner, *Tort Law: Cases and Economic Analysis* 1–9 (1982) (“This casebook is premised on the belief that the Hand formula—more broadly, economic analysis—provides a unifying perspective in which to view all of tort law.”); Richard A. Epstein, *Cases and Materials on Torts* 191 (7th ed. 2000) (“Although Learned Hand disclaimed the formulation of any ‘general rule’ for negligence, his decision in *Carroll Towing* has spawned a burgeoning

formula—the cost of avoidance and the probability and value of the loss—are so complex and difficult to ascertain that the resulting uncertainty makes the *Carroll Towing* formula resemble a standard in practice.²⁰²

Choosing between rules and standards is usually a matter of balancing error costs (over- and under-inclusiveness in rules) with administrative costs (lack of *ex ante* notice and difficulty of *ex post* application in standards).²⁰³ If we could ignore administrative costs by assuming unlimited resources, then we should always use a standard—which achieves perfect outcomes in a zero-cost world. Moreover, in the utopian zero-cost world, not only will judges always *reach* the right outcome under a standard, lawyers can *predict* that outcome *ex ante*, since they too must have unlimited resources and zero cost. In this way, the problem of *ex ante* notice is directly correlated with the difficulty of *ex post* application.

We often prefer rules to standards in real life, because the utopian world of zero administrative cost does not exist. In a world of finite resources and non-zero administrative cost, rules are simpler to administer; and their simplicity brings greater clarity and predictability (both in *ex post* application and thus *ex ante* notice)

academic literature on the economic interpretation of negligence and, by implication, the entire tort law.”).

²⁰² See *Moisan v. Loftus*, 178 F.2d 148, 149 (2d Cir. 1949) (“It is indeed possible to state an equation for negligence But . . . injuries are always a variable within limits, which do not admit of even approximate ascertainment; and, although probability might theoretically be estimated, if any statistics were available, they never are.”).

²⁰³ Richard A. Posner, *Employment Discrimination: Age Discrimination and Sexual Harassment*, 19 Int’l Rev. L. & Econ. 421 (1999) (“Rules have higher error costs but lower administrative costs; standards have lower error costs but higher administrative costs. The relative size of the two types of cost will determine the efficient choice between the alternative methods of regulation in particular settings.”). Another reason to choose rules over standards is to cabin judicial discretion—if we cannot trust judges to exercise discretion in a manner consonant with the normative values of the broader electorate. See Craig Allen Nard, *Legal Forms and the Common Law of Patents*, 90 B.U. L. Rev. 51, 97 (2009). This does not apply in patent law. The normative framework of economics-based utilitarianism is almost universally shared in patent law. Yochai Benkler, *Siren Songs and Amish Children: Autonomy, Information, and Law*, 76 N.Y.U. L. Rev. 23, 59 (2001); see *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 480-81 (1974) (“The stated objective of the Constitution . . . is to ‘promote the Progress of Science and useful Arts.’ The patent laws . . . have a positive effect on society . . . by way of increased employment and better lives for our citizens.”).

compared to standards. Thus, a rule is preferable if the administrative cost of *ex post* enforcement of a comparable standard is greater than the error costs from over- and under-inclusiveness.

B. *Layers of Patent Scrutiny*

When viewed from the prism of rules and standards, those who make the argument that we should abandon categorical exclusion in favor of more individualized inquiries such as non-obviousness are often implicitly assuming zero administrative costs. With administrative cost ignored, *of course* we would prefer more individualized scrutiny over categorical rules; or, to use Michael Risch’s colorful words, to “apply rigorous standards of novelty, nonobviousness, utility, and specification with a scalpel rather than simply eliminating broad swaths of innovation with a machete.”²⁰⁴ The problem is that the more fine-tuned and individualized the analysis becomes, the more expensive it will also be.

This becomes more evident if we consider the structure of patent law as a whole. At bottom, the policy goal of the patent system is a standard: give the inventor the amount of reward that optimally balances the monopoly reward with the incentive to innovate. Moreover, we can achieve this policy by abolishing all current patent laws except one: the provision of Section 284 that damages for infringement shall comprise a “reasonable royalty” that is “adequate to compensate for the infringement.”²⁰⁵

In a perfect world of zero administrative costs, no other mechanism is needed to achieve perfect results. If an inventor applies for a patent on an obvious invention, we should grant the patent; and then assess the “reasonable royalty” as a mere few pennies. After all, an “obvious” innovation still creates some technological advancement and social benefit, albeit very slightly. And if we had infinitely fine calibration of damages, then we can and should make the reward proportionate to that very slight contribution. This surely is better than the over-inclusive rule that all obvious inventions receive no patent and thus no reward at all, when they should optimally receive a small reward.

In the same vein, we could grant patents of unlimited term and without scope restriction, if we then tinker with the damages formula. Thus, the inventor of the paper airplane could have a perpetual patent covering “all flight devices”; but when he sues Boeing, courts would then assess the “reasonable” royalty that is “adequate” to compensate

²⁰⁴ Risch, *supra* at note 1, at 658.

²⁰⁵ 35 U.S.C. § 284 (2006) (damages for infringement).

as a fraction of a cent. This, too, allows the patent system to operate via “scalpels” rather than “machetes.”

The problem with using damages assessment as an infinitely malleable tool of precise tailoring, of course, is that the inquiry is frightfully difficult as a matter of judicial administration. Courts simply have no idea how to determine what is a “reasonable” royalty and what is “adequate” to compensate for infringement.²⁰⁶ Thus, superimposed above the damages inquiry are various “rules” that are somewhat less precisely tailored, but are easier to administer in practice. Courts preemptively decree that all obvious inventions shall receive no patent and no royalty at all, thereby avoiding the difficult question of determining a “reasonable” royalty in such cases. The loss of incentive for some trivial advances—which in a perfect world we would have wished to maintain—is outweighed by the saving in administrative cost.

The label of “rules” and “standards” here is relative. Obviousness is more rule-like and categorical than a regime of assessing the “reasonable” royalty in a totality-of-circumstances test; but it is very individualized and standard-like compared to categorical rejections such as those of products of nature. The patent system necessarily makes trade-offs between rules and standards; and arguing that individualized scrutiny and precise tailoring is always better ignores the administrative cost advantages of rules. To assume a perfect decision-maker with unlimited knowledge and resources—who can then always precisely tailor patent rights to match contribution with reward—is to assume the basic problem of administrative government away.²⁰⁷

Rather, layers of screening makes sense in order to reserve detailed scrutiny for a subset of cases. Just as it would be administratively impossible for employers to give every job applicant a callback interview, and impossible for journal editors to give every article a full committee read, it is administratively inefficient for the PTO to give every patent application a detailed individualized examination. The consequence of attempting to give detailed consideration to each of the

²⁰⁶ See John M. Golden, *Principles for Patent Remedies*, 88 Tex. L. Rev. 505, 508 (2010) (“We really have little specific sense of what the value of remedies for patent infringement generally is or should be. And it seems unlikely that we will develop a precise idea anytime soon.”).

²⁰⁷ See F. A. Hayek, *The Use of Knowledge in Society*, 35 Am. Econ. Rev. 519 (1945) (describing the “knowledge problem”).

nearly 500,000 patent applications filed annually,²⁰⁸ not surprisingly, is the PTO's large and constantly growing backlog.²⁰⁹

Of course, another solution to the scarcity of administrative resources is to increase them, such as by hiring more patent examiners. But administrative resources do not come out of thin air, and so this option is not cost-free. In the case of the PTO, the cost hiring additional examiners is passed on to patent applicants through higher filing fees.²¹⁰ Higher filing fees, in turn, reduce the patent incentive since they are functionally a tax on patent rewards.²¹¹ Therefore, too much individualized scrutiny can be as harmful to innovation as too little. Efficiently allocating administrative resources—including by using over-inclusive categorical rules to save administrative cost—is important to optimizing the patent system as a whole.

Patent law in fact employs the layers of screening approach. At the first layer, the PTO applies categorical subject-matter exclusions such as rules against patenting natural products, atomic weapons, and human clones. Although subject to collateral review during litigation, the ease of applying these limits means that such litigation is practically non-existent.

At the second layer, the individualized doctrines of novelty, usefulness and non-obviousness, as well as scope limitations such as the enablement and abstract idea doctrines, are applied by the PTO. In theory, this should result in a very precisely tailored patent. In practice, however, the PTO simply does not have the administrative resources to conduct a perfect examination. Thus, the individualized examination at the PTO is cursory, with low administrative cost but

²⁰⁸ United States Patent and Trademark Office, *Performance and Accountability Report: Fiscal Year 2009* 113 (2009).

²⁰⁹ *Id.* at 114 (showing backlog increasing from 414,837 to 1,207,794 over the last decade).

²¹⁰ Janice M. Mueller, *Patent Law* 24 (3d ed. 2009) (“the USPTO has been fully ‘user-fee funded’ for several years”). It is important to note that, although the PTO is self-funded, taxpayers still subsidize the administrative cost of the patent system, since patents also require taxpayer-funded courts to administer through adjudging infringement suits.

²¹¹ John R. Allison & Starling D. Hunter, *On the Feasibility of Improving Patent Quality One Technology at a Time: The Case of Business Methods*, 21 Berkeley Tech. L.J. 729, 785 (2006) (“Because the PTO is supported by user fees, any wasted resources represent a tax on innovation.”). The other alternative, to do nothing and let the backlog grow, also functions as a tax on patent incentives. This is because the patent term is calculated as 20 years *from filing*, so the time taken by examination generally eats away at the patent term and reduces patent rewards.

consequently high error cost. A much more rigorous and expensive application of these individualized criteria occurs during litigation.²¹² Since only a tiny portion of issued patents are actually valuable enough to warrant litigation and the consequent close judicial scrutiny, this institutional arrangement also conserves administrative resources.²¹³

The final layer of tailoring is the determination of damages. This, of course, occurs only for the small number of patents that result in litigation; and only if the patentee passes every other screen. Such an arrangement is normatively desirable, because the determination of what constitutes “damages adequate to compensate for the infringement” is the ultimate question of the patent system. This ultimate and most open-ended, difficult, expensive, and precisely tailored inquiry is reserved for the smallest number of cases.

One important point about the interaction of rules versus standards with layered screening is that the implementation is as important as the formal content of the doctrine. In theory, doctrines such as non-obviousness and scope delineation are standards—precisely tailored with low error cost and high administrative cost. But the administrative cost of any standard can in fact be drastically reduced, just by implementing it *badly* and on-the-cheap, thereby increasing error costs. A policy of examining patent applications cursorily reduces administrative costs, but it permits socially detrimental patents that are obvious or overbroad to issue. Given a fixed low level of administrative costs, we can either adopt a formal rule with built-in error costs, or we can adopt a formal standard and cause error through imperfect administration.

Stated another way, the benefit of threshold categorical rules is that they permit the more perfect implementation of individualized standards for the patent applications that remain. Since better scrutiny is necessarily more expensive—requiring more prior art searching, more testing, and more information about the social benefits and harms of a patent—finite administrative resources must be allocated. Holding administrative resources constant, for *some* patent applications to receive rigorous scrutiny implies that *others* must receive less scrutiny or none at all. To be equally detailed as to all comers is, in reality, to be equally cursory.

²¹² See Paul Goldstein, *Copyright, Patent, Trademark and Related State Doctrines* 384 (rev. 3d ed. 1993) (arguing that “the two tribunals are effectively serving different functions”).

²¹³ Lemley, *supra* at note 115.

Given the necessity of allocation, the benefit of section 101 categorical rejections is that the selection can at least be directed to categories that, on the whole, are socially detrimental or “bad” patents. For example, if 99% of business method patents are socially detrimental and 1% are beneficial; it is likely better to reject all 100% of business method patents. The categorical exclusion would be over-inclusive, in that in a perfect world we would like to sort out the good and bad business methods. But to devote resources to preserving the 1% of good business method patents would require giving less attention to, say, pharmaceutical patents, where the ratio between good and bad patents might be 50-50. The result would be erroneous rejection or allowance of some number of pharmaceutical patents due to the cursory examination, resulting in a far higher error cost than the 1% of business patents that are erroneously rejected.

One response might be that cursory PTO examination is a two-way street. It is as possible for the PTO to allow a patent after cursory examination as it is to reject the patent after cursory examination—both will save administrative costs. To some extent this is true, since the PTO’s cursory examination under the standards of novelty, usefulness, non-obviousness and enablement conserve administrative resources, allowing courts to apply these standards more rigorously for the small subset of issued patents that are litigated. But categorical allowance is not as effective as categorical denial, because the administrative costs of the patent system do not end with the issuance of a patent but are only transferred from the PTO to the courts. A categorical exclusion ends the question forever, but a categorical allowance of the patent leaves the potential that courts must determine the patent’s scope and assess damages for infringement, resulting in further administrative costs down the line.

The sum of the above is that formal categorical exclusions are a more principled and (ironically) better-tailored first screen than cursory individualized examination for all comers. The question is not whether categorical exclusion is desirable; but simply *how best to design* the categorical exclusions so that they focus on categories that have low social value and thus less error cost when excluded, while maintaining the low administrative costs of a categorical rule.

C. Principles for Categorical Exclusion

1. Cost-benefit balance.

The basic principle of categorical exclusion is that they should be directed to categories that are, on the whole, socially detrimental. In most existing literature, particularly those that seek to restore the categorical exclusion of business methods, this criterion takes the form

of arguing that business method patents are frequently obvious under existing law.²¹⁴ Supporters of business method patents then respond that business method patents are not actually invalid at higher rates,²¹⁵ and, in any event, invalidating these bad patents under obviousness doctrine is the better-tailored approach.²¹⁶ A similar argument-and-response cycle occurs in regard to patents on software.²¹⁷ In all cases, the basic critique of section 101 categorical exclusion is that they serve inappropriately as “proxies” for the novelty, usefulness, and obviousness inquiry.²¹⁸

Two contributions to this seemingly intractable debate can be made by adopting the perspective of rules versus standards that I have outlined in this Article. The first is that categorical exclusions are not proxies for novelty, usefulness, or obviousness. Rather, *all* of these doctrines are proxies for the same “ultimate” question of patent law: whether granting a patent for an invention would be socially beneficial in terms of providing more incentive benefits than monopoly costs. Categorical exclusions and individualized exclusions both seek to approximate this fundamental policy goal—simply with different trade-offs between error costs and administrative costs.

The implication is that whether a section 101 exclusion works should not be judged solely by whether the excluded patents would be invalidated by some other doctrine. That inquiry is sometimes helpful—if 99% of business method patents are actually obvious, that would be a strong reason to categorically exclude them and save the administrative cost—but it is not determinative. Obviousness is an under-inclusive rule: while almost every obvious patent is socially detrimental; not every socially detrimental patent is obvious. In the

²¹⁴ See, e.g., Dreyfuss, *supra* at note 20, at 267-68 (“Admittedly, what is particularly bad about these patents is that the methods they protect were well known before the applicant came along.”).

²¹⁵ John R. Allison & Emerson H. Tiller, *The Business Method Patent Myth*, 18 Berkeley Tech. L.J. 987, 1077 (2003) (“Our data . . . suggest that [business method] patents were of a quality . . . at least as high as most other patents.”).

²¹⁶ See Duffy, *supra* at note 24, at 622-23.

²¹⁷ Compare Jay P. Kesan & Andres A. Gallo, *Why “Bad” Patents Survive in the Market and How Should We Change?—The Private and Social Costs of Patents*, 55 Emory L.J. 61, 76 (2006) (“The existence of improperly granted patents is especially common in high technology sectors such as software and the Internet.”) with John R. Allison & Ronald J. Mann, *The Disputed Quality of Software Patents*, 85 Wash. U. L. Rev. 297, 333-34 (“software patents as a group compare quite favorably to patents that the same firms are obtaining, at the same time, on nonsoftware inventions”).

²¹⁸ Osenga, *supra* at note 1, at 1115-16.

case of business method patents, John Allison and Emerson Tiller have shown that business method patents have a private value comparable to other types of patents,²¹⁹ suggesting their general legal validity under current tests such as obviousness (since owning a legally invalid patent would give the patentee very little private benefit). But the strongest objection to business method patents is not that such methods are not valuable or useful, nor even that they are frequently obvious; it is that granting patents on such methods is often not necessary because pre-existing incentives of the marketplace to develop better business models already suffice.²²⁰ The sufficiency of pre-existing incentives does not find a natural doctrinal home in any of the other doctrines of patent law; but granting patents in fields that already have sufficient incentive to innovate would be wasteful and thus socially detrimental.

The second contribution to the literature is that administrative costs cannot be ignored in the discussion of costs and benefits under section 101; rather, it is fundamental to the structure of the system. It is true that the underlying policy goals served by categorical exclusion—granting only socially beneficial patents—can be served better by more individualized means, such as by scope adjustment or even damages calibration.²²¹ But the same is true of any legal regime; a standard will by definition achieve more precise outcomes if administrative costs are ignored. Thus, it is true that the PTO and the courts utilize section 101 rejections “to avoid the hard analyses required by other statutory sections,” a practice that Kristen Osenga condemns.²²² Contrary to Professor Osenga, however, the ability to avoid more administratively difficult and complex (and thus expensive) inquiries such as obviousness, scope, and damages is a feature, not a bug, of section 101 categorical rules.

An important implication of this observation is that exclusion is preferable because of its unique ability to cut-off further administrative expense down the road. This is illustrated by the alternative, which is heightened scrutiny at later stages of examination. One real life example is the PTO’s response to the elimination of the business method exclusion. After business methods became patentable subject-matter, and in response to widespread

²¹⁹ Allison & Tiller, *supra* at note 215.

²²⁰ Olson, *supra* at note 22, at 228 (“Particularly in the case of business methods, the level of incentive to invent new and useful business methods is quite high without any patent protection, and costs should be relatively low.”).

²²¹ For discussion of this argument, see *supra* Part I.B.

²²² Osenga, *supra* at note 1, at 1116.

concern about invalid business patents being issued, the PTO singled out business method patents heightened scrutiny. Thus, business method inventions are given better trained examiners, more prior art searching, and a second-level review.²²³

These “reforms” turn the administrative cost rationale for categorically excluding business methods on its head. The effect of assigning better trained examiners, and spending more time searching for prior art, is to *increase* the administrative resources devoted to business method patents. To be sure, this means that fewer old, useless, or obvious (in the narrow legal sense) business method patents get issued; and statistics suggest that this is the case.²²⁴ But the unappreciated collateral consequence is to *decrease* the resources available to *other* patent applications, which now have fewer and worse-trained examiners—at least if we assume that overall administrative resources are held constant.²²⁵ The result of more work being done by fewer people is either longer pendency times or higher error rates for these other patents. In the case of the PTO, a flood of business method patent applications have caused dramatic increases in pendency times across *all* areas.²²⁶ Whether there has also been an increase in error rates is more difficult to discern, since there is no readily available and objective measure of examiner error. But the increase in pendency, alone, imposes severe costs as PTO delay reduces the patent term and thus imposes a tax on patent rewards and innovation.²²⁷ The abolishment of the business method exclusion, and the consequent increase in administrative costs, has not been cost-free to inventors as a group.

Applying the cost-benefit principle to particular categories is an empirical exercise. Thus, decisively answering whether business methods, software, or any other category is properly excluded is beyond the scope of this Article. It remains worth noting that James Bessen

²²³ Allison & Hunter, *supra* at note 211, at 734.

²²⁴ Allison & Tiller, *supra* at note 215, at 1077.

²²⁵ And if administrative resources are to be increased, social resources from some other productive use must be diverted into patent administration, such as by means of a tax on research and development expenditures. See *supra* text accompanying notes 210–211.

²²⁶ USPTO, *supra* at note 208, at 114 (showing large increases in backlog and thus pendency); see Allison & Tiller, *supra* at note 215, at 1065 (“there have been dramatic increases in the number of patent applications in recent years, straining the PTO’s resources and resulting in longer average pendency times in all technology areas”).

²²⁷ See 35 U.S.C. § 154 (2006) (creating a patent term expiring 20 years after initial filing).

and Michael Meurer have provided strong evidence that business method patents and software patents are generally socially detrimental.²²⁸ And although John Allison, Emerson Tiller and Ronald Mann have provided evidence that business method and software patents are often *privately* valuable,²²⁹ there is an important distinction between private and social benefits. Based on the available evidence, there would be a strong case for excluding business methods and software processes from patentable subject-matter.

2. Ease of administration and definition.

The second principle consideration of defining categories of exclusion is their administrability. In order to maintain the administrative cost savings of categorical exclusion, the exclusion must be reasonably easy for PTO examiners and judges to apply. This generally requires that the category have reasonably clear borders.

On this criterion, many critics argue that the category of “business methods” is inherently difficult to define; and that, for any definition, patentees can easily redraft their claims to evade it.²³⁰ These are actually two separate objections, and will be analyzed separately below.

Initially, the category of “business methods” appears rather amorphous, in that it is susceptible to many definitions. At its narrowest, business methods are methods of transacting money. This would be a narrow definition, in that the definition does not cover methods of advertising, of corporate organization (e.g. a patent on the corporate form), nor barter transactions—things that might plausibly be considered business methods in common parlance. But the definition would also solve most of the vagueness problem. Given this, the fact that the definition would be under-inclusive is not a reason to abandon the categorical analysis entirely, since that would be even more under-inclusive, by not including anything at all.²³¹

Moreover, if we wished to also create section 101 exclusions for methods of advertising, methods of corporate organization, and

²²⁸ James Bessen & Michael J. Meurer, *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk* 153 (2008).

²²⁹ Allison & Tiller, *supra* at note 215; Allison & Mann, *supra* at note 217.

²³⁰ Allison & Hunter, *supra* at note 211, at 788-89.

²³¹ The category of business methods also has the special feature that, because Congress has enacted a special “prior user” defense to business method patents, courts cannot avoid the definitional problem. See 35 U.S.C. § 273 (2006). Even without any categorical exclusion, courts will still confront the problem of determining what constitutes a business method patent when determining liability for infringement.

methods of bartering, we could simply create additional categorical exclusions for those things, too, with their own individual definitions. Thus an amorphous “business method” exclusion can be made precise by transforming it into four separate exclusions, of “money transactions,” “barter transactions,” “methods of advertising” and “methods of organization.” As Kaplow observed, greater precision in a rule can always be achieved by enacting a set of multiple rules with greater complexity, though also at greater cost in rule-formulation.²³²

The more forceful objection is that, the more rigidly a category is defined, the easier it would be to evade through clever claim drafting. For example, if methods of transacting with money were deemed categorically unpatentable, Amazon.com would try to evade such a restriction on the One-click patent by rephrasing its claim as: “A *system of computers* that run a program wherein consumers can make purchases with a single click.”²³³ The addition of the generic computer element transforms the claim from one directed to a business transaction to a claim directed to a network of computers—and computers are clearly patentable subject-matter under any reasonable conception.

It is important to note that strategic claim drafting does not exploit the vagueness of category definitions, which is the first objection addressed above. The strategic drafting technique works no matter how broadly or narrowly a “business method” is defined—since under no possible definition of the term “business method” does it cover a network of computers. Instead, the technique works because, under the Supreme Court decision in *Diamond v. Diehr*,²³⁴ a court engaging in section 101 analysis must consider the “claim as a whole.”²³⁵ Because a court must consider the claim in its entirety, the addition of a known and widely used “computer” to an otherwise unpatentable business method transforms the claim into one on a computer. Patentees then receive the benefit of a Texas two-step: the invention is a computer for section 101 purposes, and therefore patentable subject-matter; but for novelty and non-obviousness purposes, the invention is the novel business method being implemented by the computer, and thus passes those criteria too. By using this Texas two-step, any

²³² Kaplow, *supra* at note 94, at 286-88 (making the comparison between a simple rule and a rule with the same complexity as a standard).

²³³ Allison & Hunter, *supra* at note 211, at 736 (“attorneys had little difficulty drafting patent applications on software as though they claimed machines and devices of a more traditional physical nature”).

²³⁴ 450 U.S. 175 (1981).

²³⁵ *Id.* at 188.

section 101 categorical exclusion can easily be evaded. All a patentee need do is add some known and widely used elements (most particularly a computer) to transform the unpatentable into patentable subject matter.

Instead, the analysis should focus on the novel and non-obvious aspects of a patentable embodiment, and exclude insignificant additions. This is known as the point-of-novelty approach,²³⁶ and is part of current doctrine, but only for one type of subject-matter exclusion: printed matter.²³⁷ Expanding the point-of-novelty approach will make many other section 101 categories meaningful and more administrable.

The printed matter exclusion means that books, videos, and CDs are not considered patentable subject-matter.²³⁸ The exclusion has teeth because courts employ a point-of-novelty analysis. Under a “claim as a whole” approach, any patentee with a novel story could write a claim in this format: “A composition of matter comprising paper, with the text XYZ.” Applying the *Diehr* two-step, the patentee could then argue that that paper—being a “composition of matter”—is patentable subject-matter, while the printed text inside is new and non-obvious. By employing a point-of-novelty analysis, however, courts would instead reject the patent.²³⁹

Two objections can be raised to the point of novelty approach. The first is that the point of novelty approach invalidates every patent. Because all inventions manipulate natural matter according to natural processes, the point of novelty of every invention can be characterized, at some level, as a natural phenomenon or natural product.²⁴⁰ But while this objection has force when applied to the products of nature and natural phenomenon exclusions, it is not true as applied to any other exclusion. Not every invention ultimately reduces to a business method, or software, or any other category that is commonly proposed for exclusion from patents. Nor must all of section 101 be subject to a single approach: as demonstrated by the printed matter exclusion, it currently is not subject to such a unified approach.

²³⁶ See Kevin Emerson Collins, *Propertizing Thought*, 60 SMU L. Rev. 317, 356-57 (2007) (arguing for a point of novelty approach to discerning mental process patents).

²³⁷ *In re Ngai*, 367 F.3d 1336, 1337-38 (Fed. Cir. 2004).

²³⁸ *In re Sterling*, 70 F.2d 910, 912 (C.C.P.A. 1934).

²³⁹ *Nagi*, 367 F.3d at 1339.

²⁴⁰ *Funk Brothers Seed Co. v. Kalo Inoculant Co.*, 333 U.S. 127, 134-35 (1948) (Frankfurter, J., concurring).

The second objection is that, by requiring an analysis of the point of novelty, I am undermining the very administrative cost saving that categorical exclusion is designed to foster. To analyze the point of novelty, one must necessarily discern what is *not* novel, and this would require searching the prior art. The entire point of categorical exclusion, however, is to avoid such a prior art search in the first place.

This would be a concern if the point of novelty inquiry required actual prior art searching to discern that the insignificant element is not novel. But the criticism misses how the *Diehr* two-step works in practice. In order for the trick to work, the insignificant element added must be universally known and in widespread use—such as by adding generic computers or the internet. If the patentee adds even a semi-obscure element that a moderately educated patent examiner might not know off-the-top-of-his-head, such as a rare type of computer, then the patent's scope would be reduced to almost nothingness—very few people could infringe the claim because only those few people use the type of computer required. In short, no patentee would attempt such a futile gesture in the first place, absolving the PTO of the necessity of defending against such a tactic. Even if the PTO adopted a policy of never conducting prior art searches in administering section 101, adopting a point of novelty approach would still provide plenty of teeth to categorical exclusions such as one for business methods, while avoiding the administrative costs of individually examining these generally meritless patents.

CONCLUSION

For clarity of analysis, section 101 must be understood to have separate components of categorical exclusions and scope limitations. These two types of doctrines have fundamentally different attributes on the rules-standards continuum. The mixing of these two types of doctrines under the rubric of a unified section 101 analysis has confused both the academic discussion and judicial analysis.

Once categorical exclusions and scope limitations are separated out, the theoretical objections to section 101 restrictions are severely undermined, leaving only empirical questions of cost-benefit balance. Categorical exclusions are sometimes over-inclusive, but have low administrative costs. Individualized examination and scope delineation is precisely-tailored, but only if given sufficient administrative resources, which necessarily requires starving other patent applications of those same resources and individualized attention. The question is thus not *whether* we should have any categorical exclusions, but one of *which* categorical exclusions are and are not empirically justified. Such empirically-based analysis would

have the beneficial effect of bring more individualized scholarly analysis to section 101, in place of overbroad calls to eliminate categorical exclusions entirely.