

Note to IPSC Readers:

As will be obvious to you if you have a chance to read it, this draft is devoid of footnotes. I have benefitted from a lot of prior work in writing this (perhaps yours!) and I will, of course, be acknowledging that in later drafts.

Thank you in advance for your feedback. I look forward to seeing you soon!

Kathy Strandburg

Patentable Subject Matter from First Principles
Katherine J. Strandburg¹
New York University School of Law

I. Introduction

The patentable subject matter doctrine is the law's way of delineating the type of thing that cannot be claimed in a patent, no matter how new or useful it might be. Since at least the middle of the 19th century, the Supreme Court has held that patents may not be awarded to natural phenomena ("laws of nature"), abstract ideas, and products of nature. In part because of advances in biotechnology and information technology, recent years have seen a resurgence of interest in the patent eligible subject matter question by the Supreme Court, which has decided four cases on the doctrine since 2010. Despite the attention the patentable subject matter doctrine has received from the Court and commentators, it remains the subject of a general malaise, not only among those who disagree with the thrust of the Supreme Court's rulings, which has been to rein patentability in, but also among many who are in substantial agreement with the outcomes in these cases.

In my view, the widespread malaise with the doctrine stems from its lack of clear theoretical bases. As Rebecca Eisenberg noted several years ago, "patentable subject matter doctrine suffers from a lack of clarity not only as to what the applicable rules are, but also as to what those rules are supposed to accomplish." Despite the recent surge in attention to the doctrine, courts and even commentators often remain mired in a sometimes scholastic exercise of interpreting the traditional exclusions of abstract ideas, natural phenomena, and products of nature, without interrogating their underlying justifications. When the discussion focuses on the goals of the doctrine, it often attempts to derive those goals from the traditional exclusions themselves, leading to a somewhat circular debate.

This Article attempts to engage the patentable subject matter doctrine directly from first principles by asking the foundational question: What types of inventions (or creative outputs more broadly) should be patentable? I take a broadly utilitarian perspective on this question here, though other approaches to the patentable subject matter question certainly are possible. Utilitarian analysis is the primary driver of patent theory and doctrine. Both courts and commentators routinely ground patent law's bread and butter nonobviousness and scope doctrines in an attempt to balance the benefits of patent incentives against the higher prices and drag on downstream innovation inevitably produced by exclusive rights in inventions. Indeed, many argue that, if these doctrines are properly tailored, there is no (or extremely limited) need for patentable subject matter exclusions. I argue here that the patentable subject matter doctrine serves important purposes that cannot be addressed by case-by-case analysis of nonobviousness and claim scope.

The purpose of the patent system is to solve certain market failures and thus incentivize the invention, disclosure and dissemination of innovative technology. The patent system may not always be the most socially attractive alternative, however. In some arenas, where the market failures are minor or the costs of patenting are particularly high, the patent-free market may be

¹ Alfred B. Engelberg Professor of Law, NYU School of Law, katherine.strandburg@nyu.edu

socially preferable to a patent-based market. Many critiques of business method or software patenting essentially make this argument. Patents also are aimed only at certain types of market failure. They are ineffective if market demand fails to reflect social demand. Market demand famously fails at providing infrastructure, due to collective action problems and free riding. It also fails to account for both positive and negative externalities of consumer's individual purchases. Market demand reflects ability to pay, which may or may not accord with social value. It tends to be myopic, meaning that investments in innovation are likely to be both too small and misdirected from a long-term perspective. Finally, there are a number of alternative approaches to correcting innovation-related market failures, ranging from government subsidy to informal norms. Some of these alternatives, such as Internet-based crowdsourcing, have emerged out of the digital communications revolution, while others, such as the norms of scientific research, have deep historical roots. Depending on the technological context, these alternative institutions may be more or less effective than patents, may be more or less costly to administer and use, and may have differing implications for the *direction* of innovation and the distribution of its benefits.

The doctrines of nonobviousness and claim scope sit squarely within the patent-based market paradigm, addressing the question of which exclusive rights should be granted within a patent-based market system. The patentable subject matter doctrine should take on a different task. It is the doctrinal mechanism for determining *whether a patent-based market is the best system for promoting innovation of a particular type*. Unfortunately, patentable subject matter doctrine has not, for the most part, grappled with this question. The doctrine's failure to confront the systemic question directly might be of little consequence if alternative approaches simply operated in parallel with the patent-based market. In general, however, the availability of patents can be expected to influence the way in which alternative innovation institutions operate -- and sometimes even to undermine them. Thus, patentable subject matter doctrine inescapably molds institutional choices, whether or not we choose to recognize its effects. Currently, it does so inconsistently and obliquely. That is a mistake. A more forthright recognition of the doctrine's institutional implications would sharpen and clarify the doctrine and perhaps provide a partial cure to its longstanding malaise. Managing the interface between the patent system and alternative innovation institutions should be the primary job of the patentable subject matter doctrine, to be carried out for the most part through categorical exclusions.

Recognizing that there alternatives to the patent system in some contexts puts the patentable subject matter doctrine in a new light and raises several compelling questions: Under what conditions should we expect that an alternative institution will outperform the patent system? When can such systems co-exist happily with the patent-based market and when should we use patentable subject matter exclusions to open up space for them? How would we redesign patentable subject matter doctrine to account for institutional and other non-market considerations? This Article cannot hope to provide ultimate answers to these questions, but it attempts to make a start at addressing them.

Part II of this Article provides background for the patentable subject matter discussion by explaining how the patent-based market paradigm is enshrined in basic patent doctrines. Part III argues that patentable subject matter doctrine flounders when it stays within that paradigm. Part IV describes examples of alternative innovation institutions and considers the circumstances

under which patentable subject matter exclusions may be appropriate to sustain them. Part V provides a tentative outline of an institution-based approach to patentable subject matter doctrine. Part VI explores, as an example, how such an approach to patentable subject matter would play out for natural phenomena, arguing that the Supreme Court’s opinion in *Mayo v. Prometheus* comes close to applying an institutional analysis. Part VII concludes.

II. Market Improvement: Nonobviousness, Scope and the Patent-Based Market Paradigm

The Constitutional purpose of the patent system is to promote progress in the useful arts. Patents accomplish this goal by, as Abraham Lincoln evocatively put it, “adding the fuel of interest to the fire of genius.” Without patents, Lincoln explained, “any man might instantly use what another had invented; so that the inventor had no special advantage from his own invention.” To put it in more modern terms, the patent system addresses a market failure due to the fact that some inventions can be copied by “free riding” competitors before sufficient payoffs accrue to cover their costs of invention. Patent exclusive rights ensure higher market returns, thereby incentivizing invention. Patents also address two other types of market failure. Some inventions can be exploited in secret. While secrecy solves the free riding competitor problem, it can impede downstream innovation that would build on the new invention. Patent exclusivity provides a “quid pro quo” incentive for disclosure. In addition, Arrow’s paradox suggests that cooperation between inventor and potential funders or commercializers will be deterred because inventors will be reluctant to share their ideas with potential partners who might simply use them without paying. By assigning rights in advance of such discussions, patents may facilitate commercialization.

Patent exclusivity is socially costly, however, because it constrains the implementation of ideas, which optimally, to quote Thomas Jefferson, “should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition.” Patents lengthen the “first mover advantage” period, during which inventors can restrict output and charge supracompetitive prices, increasing deadweight economic losses, costs of downstream innovation, and redistribution from consumers to patentees. Even when an inventor chooses patenting over secrecy, a patent ordinarily will extend the market exclusivity period, imposing similar social costs.

Patents, to summarize, are a necessary evil intended to overcome particular market failures. From its inception, U.S. patent law has aimed to balance the patent system’s costs and benefits by providing exclusive rights only when the benefits in resolving market failures outweigh the social costs of those rights. Patent doctrine seeks to achieve this balance in two primary ways: i) the doctrines of novelty and nonobviousness deny patents for insignificant advances and ii) a set of scope doctrines attempt to match exclusivity payoffs roughly to patentees’ inventive contributions. This Part briefly describes how these doctrines aim to achieve this balance.

A. Nonobviousness and the Realistic Marketplace

The doctrines of novelty and nonobviousness determine whether a patent applicant has done “enough” to warrant patent rights, as evaluated from the perspective of the “person having ordinary skill in the art,” or PHOSITA. The economic rationale for this set of requirements is

that first mover advantages in the competitive market are sufficient to induce the modest investments presumably required to make inventions that would be “obvious” to the PHOSITA. Because of their social costs, patent rights should not be granted for these “obvious” inventions.

Within this general picture, different interpretations of the nonobviousness standard reflect different views about how much innovation the competitive market will produce on its own. For example, the Supreme Court’s most recent foray into the nonobviousness doctrine, *KSR v. Teleflex*, overturned the Federal Circuit’s approach primarily because of a disagreement about what the PHOSITA can be expected to produce without the expectation of a patent. The Federal Circuit had held that a claimed invention could be deemed obvious only if the state of the art provided a “teaching, suggestion, or motivation to combine” existing prior art to produce that invention. The Federal Circuit’s test thus reflected the view that, without patents, market actors would do very little beyond what was spelled out in the prior art. In over-ruling the Federal Circuit’s approach, the Court faulted the Federal Circuit’s unrealistic assessment of the power of market competition to induce innovation. The Court noted that “[w]hen a work is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or a different one.” Similarly, “[w]hen there is a design need or market pressure to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within his or her technical grasp.” The Court also observed that “[a] person of ordinary skill is also a person of ordinary creativity, not an automaton” and thus the analysis should “take account of the inferences and creative steps that a person of ordinary skill in the art would employ.” The Court was concerned that the Federal Circuit’s approach awarded patents on the results of “ordinary innovation” that would have been induced by market competition, thus imposing social costs that “might stifle, rather than promote, the progress of useful arts.”

Though *KSR* incorporated a more dynamic view of the competitive market, it did not step outside of the market paradigm. Moreover, it did not account for the way in which non-market forces, such as intrinsic motivations or the desire to use an invention, might reduce the market payoff required to induce a particular investment of inventive effort. In such cases, a stiffer nonobviousness requirement might be appropriate. So far, the nonobviousness doctrine, even when adapted to a more realistic view of the competitive market, has not accounted for non-market motivations.

B. Scope Doctrines from a Market Perspective

A number of patent doctrines, such as enablement, written description, claim overbreadth and indefiniteness, work to tailor claim scope to what the patentee actually invented. Assuming a patent is granted, claim scope, coupled with market demand, determines the size of a patentee’s payoff from her exclusive rights. Matching claim scope to a patentee’s actual invention means that the exclusivity payoff is correlated to the “size” of the invention. This strategy is sensible if one assumes that, at least on average, inventors must make larger investments to develop “bigger” inventions. Confining claim scope to the patentee’s invention thus provides an approximate means for matching exclusivity payoff to inventive investments. Claim scope doctrines aim to ensure inventors receive sufficient, but not excessive, returns, thus avoiding

unnecessary social costs from overly broad patent rights. These doctrines – particularly the indefiniteness doctrine -- also attempt to reduce transaction costs by ensuring that claim boundaries are well-defined.

To sum up, while there are ongoing debates about their precise formulations and parameters, most substantive patent law doctrines attempt, in some reasonably sensible way, to resolve particular innovation market failures while balancing the social costs and benefits of awarding exclusive rights to a particular patent claim. The balancing of costs and benefits implicit in these doctrines takes place entirely within the patent-based market paradigm. This discussion sets the stage for our consideration of patentable subject matter because some commentators have argued that, if properly tailored, these doctrines essentially cover the field, leaving no need for patentable subject matter exclusions.

III. One of These Things is Not Like the Others: Patentable Subject Matter Doctrine and the Market Paradigm

Patentable subject matter doctrine is different. Attempts to make sense of the doctrine from within the market paradigm have been notably unsatisfactory, leaving patentable subject matter exclusions vulnerable to critique. Discussions of the doctrine's rationale in both case law and commentary often seem unmoored from the patent system's basically down-to-earth utilitarian approach. This Part argues that the difficulty in pointing to a convincing rationale for patentable subject matter doctrine stems from the simple fact that in some respects the critics are right. While there is a sliver of patentable subject matter doctrine that makes market sense, for the most part patentable subject matter exclusions are not sensible from within the market paradigm.

A. Two Doctrinal Threads of Patentable Subject Matter Doctrine

This Article picks up in some respects from where my earlier article, *Much Ado About Preemption*, left off. *Much Ado About Preemption* pointed out that patentable subject matter discourse generally has conflated two distinct threads of analysis. According to the first, patentable subject matter exclusions are intended to reject claims that would have overly broad downstream impact or, as it is often put, “preempt” downstream innovation. The second thread is based on a determination that some category of inventions patent ineligible per se. To analyze whether a particular claim is patent eligible, one first determine whether the claimed invention incorporates an excluded element, and then applies some rule for determining whether the claim incorporates the unpatentable element in a patent eligible way. To be normatively satisfying, the second-stage rule should bear a logical relationship to the justification for the categorical exclusion.

Traditionally, the Court has held that natural phenomena, abstract ideas, and products of nature are ineligible for patent protection. While the Court often (and confusingly) employs “preemption” rhetoric in its opinions, in the majority of cases it has taken a categorical approach to those traditional exclusions. Indeed, the Court recently affirmed that evaluating patentable subject matter requires a two-step approach along the lines discussed above. Unfortunately, and confusingly, the opinions nonetheless continue to mix “preemption” rhetoric into analysis of categorical exclusions.

1. The Broad Downstream Impact Thread

O'Reilly v. Morse is the granddaddy of the broad downstream impact thread. There, the Court invalidated a very broad claim in Morse's telegraph patent. In the claim, Morse "d[id] not propose to limit [himself] to the specific machinery ... described in the foregoing specification and claims; the essence of my invention being the use of the motive power of ... electromagnetism, however developed, for making or printing intelligible characters, signs or letters at any distances" In invalidating the claim, the Court emphasized its potential impact on downstream innovation:

"It is impossible to misunderstand the extent of this claim. He claims the exclusive right to every improvement where the motive power is the electric or galvanic current, and the result is the marking or printing intelligible characters, signs, or letters at a distance. If this claim can be maintained, it matters not by what process or machinery the result is accomplished. 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."

Modern scope doctrines have ameliorated many of the concerns articulated in Morse. Today, Morse's claim would probably be invalidated as encompassing more than he actually invented. Perhaps as a result of strengthening the scope doctrines, modern patentable subject matter cases that turn on a claim's broad downstream impact are relatively rare, despite the continuing references to "preemption." The patentable subject matter determination in *Bilski v. Kappos*, decided in 2010, turned, at least partly, on broad downstream impact. *Bilski* affirmed the PTO's patentable subject matter rejection of a set of claims involving processes for hedging risk in commodities trading. *Bilski*'s discussion of the claims is fairly cursory, but the unpatentability of the patent's broadest claims seems to have been based on their broad sweep. *Gottschalk v. Benson*, holding that claims involving a binary-decimal conversion algorithm were unpatentable, also relied on the potential downstream impact of the claims. In *Benson*, however, the broad downstream impact was due to the wide variety of downstream *uses* for the claimed algorithm in the emerging field of digital computing, rather than to the breadth of the claim *per se*. As Dreyfuss and Evans have argued in the context of gene patents, a claim may affect a broad swath of downstream inventive activity even without encompassing a broad scope of embodiments. Indeed, "even very narrow patents [may] block off too much" when they are difficult to design around. In such cases, the standard scope doctrines cannot confine the downstream impact of the claims. A patentable subject matter exclusion may be used for the same purpose.

2. The Categorical Exclusion Thread

Despite its continuing use of preemption rhetoric, the Supreme Court has based nearly all of its patentable subject matter opinions on a two-step categorical analysis. *Mayo* and *Alice*, two of the Court's most recent patentable subject matter opinions, set forth a two-step analysis for the

natural phenomena and abstract ideas exclusions. In *Mayo*, the Court first determined that correlations between the blood level of a drug metabolite and the drug’s toxicity and efficacy were unpatentable natural phenomena. It then asked whether the claims reflected an “inventive concept” sufficient to ‘transform’ the [natural phenomenon] into a patent-eligible invention.” In *Alice* the Court determined, at the first step, that the claims incorporated the unpatentable abstract idea of “intermediated settlement.” It then applied the second-stage “inventive concept” rule to conclude that the claims, which did “no more than require a generic computer to perform generic computer functions” in conducting intermediated settlement, were unpatentable.

In *AMP*, which dealt with the patent eligibility of claims to DNA sequences, the Court held that an isolated DNA molecule encoding a native sequence was an unpatentable “product of nature,” while a cDNA molecule based on the same sequence was patentable because differences in the non-coding regions of the molecule make cDNA “distinct from the DNA from which it was derived,” such that “the lab technician unquestionably creates something new when cDNA is made.” Here again the Court applies a two-step categorical exclusion analysis, though it is not entirely clear what rule the Court used to draw the line between isolated DNA and cDNA or why it did not apply (or even mention) the inventive concept rule.

The opinions in the categorical exclusion thread leave a number of questions open. Despite their pedigree, the traditionally excluded categories -- natural phenomena, abstract ideas and products of nature – remain ill-defined and unclearly justified. Moreover, the Court’s rationale for applying particular second stage rules to particular unpatentable categories is still opaque. Finally, the Court continues to muddy the waters by employing preemption rhetoric in these cases.

B. Patentable Subject Matter Doctrine and the Market Paradigm

The Supreme Court’s decisions put forward a number of rationales for rejecting claims as unpatentable subject matter. As noted, the Court often speaks of the danger of “preemption” of downstream innovation by patenting of abstract ideas or natural phenomena. Often, the Court states that natural phenomena and abstract ideas should not be patented because they are “building blocks” or “basic tools of scientific and technological work.” According to the Court, “there is a danger that the grant of patents that tie up their use will inhibit future innovation premised upon them, a danger that becomes acute when a patented process amounts to no more than an instruction to ‘apply the natural law,’ or otherwise forecloses more future invention than the underlying discovery could reasonably justify.” Yet it is not immediately evident either why a patentable subject matter exclusion is needed to avoid these problems or how the two-step categorical exclusion approach implements these policy goals. This section considers those questions in turn from a market perspective.

1. Is There a Market-Based Justification for Excluding Claims With Broad Downstream Impact?

The Court’s rhetoric focuses on the impact of patents on downstream innovation. All patents make downstream innovation more costly, however, at least to some degree. How should we determine whether a claim “preempts” downstream innovation to an inappropriate extent? From a market perspective, there are three possibilities: First, a claim might have an “overly broad”

impact because the patentee receives an exclusivity payoff that is too high relative to the necessary inventive investment, meaning that it imposes unnecessarily high deadweight losses. Second, a claim might impose particularly high transaction costs. Third, it might be that the inventor of a claim is unable or unwilling, for some reason, to exploit the claim for its full social value.

a. Overcompensating the Patentee

When might a broad impact claim that is properly tailored to a patentee's invention overcompensate the patentee? Perhaps the claim covers a "problem" rather than a particular "solution." If problem finding tends to add little to the total cost of finding a solution, such claims would systematically overcompensate patentees. Patent doctrine already anticipates this possibility however, and addresses it with the doctrine of utility, which enforces the rule that "a patent is not a hunting license." Perhaps the broad impact invention was a low cost generalization of a narrower invention. Morse's broad telegraph claim, for example seems to have been based on a "cheap" generalization of his invention, which required little effort over and above the investment required for the narrower claims. Dreyfuss and Evans provide another plausible scenario: a patentee might get a windfall recovery from a patent on a relatively cheap invention that is hard to design around. As already mentioned, patent law's scope doctrines do not catch this type of overcompensation.

b. High Transaction Costs

Patent claims with broad downstream impact are likely to generate a large number of licenses, involve many licensees, and to entail complicated interactions between the various licenses. As a result, licensing costs may rise more than proportionally to the number of licenses required, meaning that such claims may tend to have especially high transaction costs. The standard nonobviousness and scope analyses do not account for this type of transaction cost concern.

c. Limited Patentee Capacity

Perhaps inventors of claims with broad downstream impact are especially likely to lack the range of information or expertise needed to manage their patent "prospects" well. This question effectively recapitulates the old debate between Kitch's prospect theory and Merges and Nelson's argument that more innovation will be produced if patents are confined to narrower prospects. While that debate certainly has not been resolved definitively, the weight of scholarly view stands behind Merges and Nelson. Once again, the standard nonobviousness and scope analyses do not take this concern into account.

It is plausible, then, that claims with broad downstream impact tend to be especially socially costly even when they comply with patent scope doctrines. Courts spend very little time, however, in assessing whether particular claims are likely to have problematically broad downstream impacts. Instead, they employ the categorical exclusion approach. But is there any reason to expect that properly scoped claims to products of nature, natural phenomena, and abstract ideas are more likely than claims to other inventions to "preempt" an unacceptable swath of downstream innovation? While some exemplars of these categories seem to fit the bill, others

do not. Moreover, there surely are inventions outside of these categories with equally broad implications for downstream innovation, yet they are never considered as candidates for patentable subject matter exclusion.

The relationships between the articulated rationales for the traditionally excluded categories and the second-step rules are also obscure. Having justified the traditional exclusions on preemption grounds, the Court is surprisingly likely to ignore the extent to which additional elements narrow the downstream impact of the particular claims at hand, as it did in *Bilski*, *Mayo*, and arguably even in *Benson*. Similarly, in *AMP v. Myriad*, the Court articulated a “basic tools” rationale for holding that native DNA is an unpatentable product of nature, but then deemed cDNA, which is nothing if not a basic tool, patentable simply because its structure was determined by a lab technician.

Moreover, as a long line of critics going back to the dissenters in *Morse* have pointed out, banning broad downstream impact claims brings us back to square one with regard to the standard market failures that patents are intended to solve. Yet the Court is singularly unconcerned with how unpatentable subject matter inventions will be produced. Justice Breyer’s dissent in *LabCorp*, for example, which was the precursor to the unanimous opinions in *Mayo* and *Alice*, illustrates this attitude:

The justification for the principle does not lie in any claim that "laws of nature" are obvious, or that their discovery is easy, or that they are not useful. To the contrary, research into such matters may be costly and time-consuming; monetary incentives may matter; and the fruits of those incentives and that research may prove of great benefit to the human race.

A rational theory of patentable subject matter must address this issue. Though natural phenomena and products of nature may be, as Judge Rader put it in his *Bilski* dissent at the Federal Circuit, provided by “God or Allah or Jahveh or Vishnu or the Great Spirit . . . as humanity’s common heritage,” the fact remains that someone has to invest in their discovery, disclosure, and dissemination if they are to be turned into socially valuable technology. Abstract ideas, similarly, are the product of human mental effort that someone must make. The patent system is grounded in the observation that markets tend to fail to induce sufficient invention, disclosure, and dissemination of technology. If not patents to solve these market failures, then what? Currently, patentable subject matter doctrine offers no answers. From within the market-based paradigm, the doctrine appears unsatisfyingly ad hoc.

IV. The View From Outside the Box: Patentable Subject Matter Doctrine as Comparative Institutional Analysis

For a doctrine designed to promote innovation, patent law is remarkably old-fashioned in its underlying model of the invention process. Elsewhere in social science and in legal discourse, scholars have spent the last thirty years or so grappling with the recognition that social institutions matter and that economic reality involves more than atomistic markets. Ronald Coase’s seminal article, the Nature of the Firm, for example, attacked the theoretical question of why economic actors organize into firms. The 2009 Nobel Prize in Economics went to Oliver

Williamson and Elinor Ostrom for their separate investigations of the theory and empirics of the social institutions that govern economic life. Ostrom, in particular, emphasized that social groups often solve social dilemmas by creating institutions to govern shared resources without resorting entirely to either government or the market. There are now many competing theories of institutional economics, emphasizing different aspects of social interaction and aimed at accounting for a variety of institutional characteristics. Patent scholars, too, have begun to analyze the institutional aspects of invention.

Despite all of this ferment, patent doctrine seems largely oblivious to the variety of institutions created to organize various aspects of social life. Its implicit model continues to assume that the options are patent system, market, or bust. This is a particularly big problem for patentable subject matter doctrine because its primary function, both descriptively and normatively, is to set boundaries between parts of the inventive terrain where creative activity is governed by the patent-based market and parts of the inventive terrain where creative activity is governed by alternative institutions. While the nonobviousness and claim scope doctrines operate on a claim-by-claim basis to fine-tune the patent system's ability to solve market failures, patentable subject matter doctrine takes patents out of the picture entirely for certain types of inventions. In these patent-free zones, other institutions unavoidably take over innovation governance. The replacement institution might be the patent-free competitive market (failures and all) or some other social arrangement, but there is no avoiding the institutional switch.

We now know that there are many patent-free mechanisms for governing and sustaining innovation. The competitive market itself is such an institution. In some fields, patents may be unnecessary because first mover advantages alone are sufficient. In others, the costs of defining and enforcing patent boundaries may outweigh the patent system's benefits. In still others, non-pecuniary motivations, such as learning, enjoyment, and use, are sufficient, either in and of themselves or in combination with first mover advantages. In some arenas non-market institutions, such as knowledge commons, govern the invention, disclosure and dissemination of some types of inventions, often combining reputational rewards with intrinsic motivations. Some of these non-market arrangements are governed by informal community norms, while others employ more formal self-governance arrangements. Groups of user innovators, for example, frequently form communities in which invention, disclosure and dissemination are fostered through a combination of use benefits and informal norms. Government subsidy is another non-patent institution for promoting and sustaining innovation. As the venerable "Republic of Science" illustrates, government subsidy can be combined in various ways with cooperative and self-governing arrangements to establishing a successful innovation system.

To provide background, this Part first describes a few alternative innovation institutions in a bit more detail. It then considers the circumstances under which such alternatives are likely to be socially preferable to the patent-based market system.

A. Alternative Innovation Systems: Some Examples

All innovation institutions have to solve several basic problems. First, they must either solve or avoid the basic market failures by providing incentives to invent in the face of free riding, incentives to disclose when secrecy is possible, and incentives to disseminate the invention to

users. Second, unless the invention can be both made and used in isolation, an innovation institution must solve the organizational issues of obtaining inputs, organizing collaborative inventive effort, and organizing dissemination of the output. The market-based patent system solves the incentive issues by providing exclusive rights and the organization issues by arms-length sales and contracts. Alternative innovation institutions address these same issues in different ways, as illustrated here by a few examples.

1. Patent-free Market Regimes

A patent-free market regime may be a viable alternative to a patent-based market in some arenas. In some cases, the traditional competitive market produces substantial innovation. In other arenas, alternative market-based mechanisms may be used to overcome the market failures that the patent system is designed to solve.

a. Traditional competitive markets

As mentioned earlier, in some arenas patents may be unneeded because there are no market failures – first mover advantages in the competitive market, alone or in combination with various intrinsic motivations, are adequate motivators. Many would argue that fashion design and business methods are such arenas, for example. In principle, the patent system should have no impact on these arenas because the nonobviousness screen would knock out patent claims where first mover advantages are sufficient. In practice, this is unlikely to happen because it would be very difficult, as a matter of political economy, for the PTO to reject essentially all patent applications in a specific arena. Once some players in a particular arena are granted patents, competitors will seek patents for defensive reasons. When such an arena can be identified, it thus is socially preferable to exclude it from patentability.

Another kind of situation in which traditional competitive markets may be preferable is when patents lead to particularly high transaction costs. High transaction costs alone are not enough to justify reliance on the patent-free market, however. There must be some reason to expect that the patent-free market will continue to innovate without patents. However, innovation is not an either/or thing. Patent-free competitive markets always produce *some* innovation. Patents, when they work as intended, simply increase the rate and amount that occurs. If, however, the transaction costs of the patent system are sufficiently high in a given arena, they may outweigh the system's benefits. Such an unfavorable tradeoff is particularly likely in arenas where first mover advantages are robust. Thus, for example, scholars have produced both argument and empirical evidence suggesting that the costs of defining patent boundaries and giving notice of patent rights are especially high in the software arena. Claiming doctrine, such as indefiniteness, may be able to reduce those costs, but if transaction costs remain particularly high for software, society may be better off doing away with software patents and relying on the traditional competitive market. This is especially likely to be the case because (at least some kinds of) software innovation occurs rapidly and is relatively inexpensive, such that first mover advantages are relatively effective means to recoup inventive investments.

b. Crowdfunding

Crowdfunding, exemplified by Kickstarter.com, is a system by which budding entrepreneurs can aggregate funding to support bringing their ideas to market. Crowdfunding generally is Internet-based. To use Kickstarter, for example, an innovator posts an idea for a product and promises to produce it if enough buyers are willing to commit upfront to pay a certain price for the completed product. Kickstarter solves the free riding competitor problem without using patents. Prospective inventors take pledges upfront, thus aggregating demand before investing in bringing a product to market. Once the inventor knows there is sufficient demand to cover the costs of designing and engineering the product, she can invest without fear of free riding. Those who have pledged in advance are contractually obligated to pay the inventor, thus preventing a competitor from undercutting the price to steal the sale. Such arrangements were rare in the past, given the high costs of finding and aggregating potential purchasers and matching them to inventors. Kickstarter lowers those costs dramatically, making such upfront agreements much more feasible. Rather than relying on patent exclusivity to recoup inventive costs, crowdfunding permits price discrimination, since inventors can tap into the higher willingness to pay of early adopters, while offering products to other consumers at lower prices.

Crowdfunding may not address all of the problems that patenting is intended to solve, however. Projects posted on a crowdfunding site must be far enough along in development to attract pledges from potential purchasers. Crowdfunding thus requires some amount of upfront investment by innovators (often the development of a prototype). Moreover, to attract supporters, an innovator must disclose a certain amount of information about the innovation, which might permit competitors to free ride on that upfront investment. However, as Michael Burstein has explained, many commercial joint ventures overcome similar problems using staged disclosure, suggesting that the disclosure needed to attract pledges may not always need to be detailed enough to allow competitors to free ride. Moreover, posting an idea on a crowdfunding site may serve as a marker of inventorship. Competitors who copy an idea posted on a widely used platform such as Kickstarter may incur reputational penalties.

Crowdfunding is an intriguing possible alternative to the patent system, but we do not yet know whether and for what types of inventions it would be a successful substitute. Currently, there is nothing to stop Kickstarter participants from using crowdfunding as a supplement to, rather than a substitute for, patenting. Indeed, some (but not all) of them do exactly that. More experience with crowdfunding (and more academic study) is needed to determine the conditions under which crowdfunding may be a feasible alternative to the patent system.

c. Advertising

Advertising is an increasingly important alternative innovation system, though it is rarely described in such terms. From its beginnings in radio and television, funding through advertising has served as an alternative to intellectual property. Intellectual property works because IP owners can charge supracompetitive prices to consumers. The broadcast media were unable to do that because they could not track which consumers consumed the programming they produced. Selling advertising space provides a stream of income that does not depend on identifying and charging consumers directly. Media companies recoup their investments from advertisers instead, who presumably pass those costs along by charging more for consumer goods. Nowadays, of course, a wide variety of businesses use advertising-based revenue models.

The advertising-based model is socially advantageous in some respects, because it does not constrain output – consumers who would not be willing (or able) to pay for an advertising-supported service obtain access to it at essentially zero monetary cost. Advertising-based business models also have disadvantages. Increasingly, advertising-based business models rely on large-scale (and often devious) collection of personal information, creating privacy concerns. They also substitute advertiser demand for consumer demand. While the two overlap (because companies do not want to buy advertising on programs that no one consumes), they are never completely aligned. This is a problem particularly, as I have argued in detail elsewhere, in the case of behaviorally targeted advertising, which skews advertiser demand toward products that collect personal information. Depending on how all of these costs and benefits net out, advertising-based business models may or may not be socially preferable to the patent system in producing particular types of innovation.

2. Non-Market-Based Innovation Systems

There are also a variety of non-market-based innovation systems. An innovation system can be non-market-based in various ways. It might rely on something other than market demand to direct the allocation of innovation resources. Such approaches are particularly valuable in arenas in which the market fails to demand what society needs and wants. Government subsidies, philanthropic sources, or contributions from the innovators themselves can supply the necessary input resources. Unlike the patent system (or any market-based system), non-market-based systems can address demand-side market failures by relying on public or philanthropic support. Alternatively, or in addition, a non-market-based system may rely on intrinsic incentives, on social norms, or on some other non-market governance mechanism to incentivize and reward innovative activity. Such systems can avoid some or all of the deadweight loss and transaction costs associated with patent exclusivity, though they will have costs of their own. This section describes some examples of non-market-based innovation systems and compares them to the patent system.

a. Prizes and Procurement

Though they have been used for centuries, prizes are experiencing a resurgence of popularity as an approach to inducing innovation. Prizes may be publicly or privately funded, but their defining feature is that an offeror defines (to at least some level of specificity) the goal to be reached or problem to be solved. Like patents, they rely on inventor self-identification and upfront investment, but, unlike patents, they do not rely on market demand. Prizes thus generally make most sense in contexts where market demand fails to match social demand. Procurement is similar to a publicly funded prize contest in that it depends on public money and top-down goal identification, but relies less on inventor self-identification and more on the government's ability to select the best inventor. Because prize and procurement systems do not rely on market exclusivity to reward inventors, they can (if appropriately designed) avoid many of the social costs of the patent system. They may, however, have significant transaction costs of their own associated with specifying the goal and awarding the prize. Moreover, if a prize system is to replace patents, it must be large enough, so that its expected value will cover the

inventive investments of all contenders. Many current prize systems do not preclude patenting and thus simply supplement the patent system's rewards.

b. Crowdsourcing

“Crowdsourcing” is a loose term used to refer to a wide variety of different innovation regimes. The common feature of crowdsourcing projects is their reliance on the way in which the Internet has drastically lowered communication costs and, as a result, the costs of seeking out those with whom it might be beneficial to collaborate. As Yochai Benkler pointed out in his seminal article on open source and “peer production,” the Internet makes it cost-effective for individuals to nominate themselves to perform tasks with and for others based on their interests and expertise.

Many crowdsourcing projects have little to do with innovation (except perhaps in the design of the project). Others, such as Innocentive, are aimed at identifying and recruiting the right person for an innovation task, generally for a financial reward. This kind of crowdsourcing operates essentially as a mechanism for improving market-based contracting for innovation. Next, there is innovative peer production, in which an innovation problem is modularized into bite-sized chunks so that the investment demanded of each contributor is small. If the chunks can be made small enough, the intrinsic benefits of participation (whether in the form of altruism, fun, a sense of civic duty, or something else) can motivate a large enough group of contributors to complete the project. Though the individual tasks are (by assumption) innovative, they cannot be too difficult or time-consuming in this model. The success of peer production as an innovation system thus depends on the care taken in designing the crowdsourcing platform and modularizing the problem. To get a truly innovative result, it seems likely that someone (or some group of someones) will have to make a substantial investment in these infrastructural design tasks. Of course, the line between sufficiently and insufficiently modularized tasks is ill defined because individuals vary greatly in their talents, skills, leisure time and intrinsic interest in and enjoyment of tasks. A half hour spent browsing the Internet demonstrates that there are people with deep interest in a surprising variety of subjects, and people willing to invest voluntarily in creating the infrastructure necessary for many kinds of crowdsourced projects.

Open source software is by far the most celebrated example of peer production. Though open source software is deployed in many commercial contexts and a large fraction of its contributors are paid to participate, I categorize it as “non-market-based” here primarily because its participants do not rely on market mechanisms to coordinate their activities. In many instances, contributors to open source software are user innovators. Given that, it is perhaps not surprising that the big success stories of open source software are infrastructural programs, such as the Linux operating system, mailing programs, and other programs that are used by large groups of programmers.

The open source software approach combines elements of the different versions of crowdsourcing discussed above. It provides a platform for collecting dispersed information from users about problems with the code. It relies on the fact that software is highly modular so that many coding tasks are small and accessible to programmers with time on their hands. It also relies, however, on the dedication of small groups of insiders who invest large amounts of time not only in coding, but also in defining issues and making decisions. These insiders often

operate within highly structured governance systems. Finally, its success depends on the significant investments of individuals who set out the initial structure and functionality of a particular program (and often retain great influence over its later development).

Open source has been wildly successful, as already mentioned, at producing infrastructural programs. It seems to have been less successful, at least anecdotally, in producing user-friendly programs for non-expert users. (Very few people run Linux on their PCs, for example.) It remains to be seen whether open source software is a viable alternative for producing software to be used by ordinary consumers. It also is unclear how far the open source model (or crowdsourced innovation more generally) can be extended. Software is a rather special innovative output, in that it can be disseminated to users almost costlessly, without the need for manufacturing facilities and the like. To the extent that production costs can be reduced and design can be separated cost-effectively from production, 3D printing permits generalization to some kinds of products. But extending crowdsourcing to less modularizable and more expensive innovation projects may be difficult.

d. Knowledge Commons Governance

“Knowledge commons governance” is a general term for innovation institutions that are neither market-based nor governed by formal hierarchy. The term is derived from an extensive literature, grounded in the work of Elinor Ostrom, which describes and analyzes the ways in which communities sustain and allocate natural resources using commons governance.

Commons governance generally means that those who work on producing a set of knowledge resources employ self-governance mechanisms to determine the direction of creative work, the assignment of tasks, and the allocation of rewards among participants. The creative work of a knowledge commons may be financed by internal resources, by public or philanthropic money, or even by market exchange with outsiders. Knowledge commons governance is often a substitute for the patent system. This section discusses some notable examples of knowledge commons.

i. Open Science

“Open science” refers to an innovation system governed (at least in the ideal case) by the so-called Mertonian norms of communalism, universalism, disinterestedness, and organized skepticism. Rewards are reputational, based on publication, and allocated by peer review. Reputational reward systems are socially advantageous, in that they do not depend on exclusivity (indeed, they encourage widespread dissemination). Reputation reward systems have their own costs, of course. Since human beings “cannot live by [reputation] alone,” reputational systems require some associated means for paying researchers and for funding research. The open science system thus is dependent on public or philanthropic funding, while scientists are supported by salaries from universities or other institutions. The open science system also avoids the transaction costs involved in determining patent claim boundaries and negotiating over licenses, though peer review and public funding introduce their own transaction costs.

The scientific research system is governed by social norms that require publication of results and, in some cases, sharing of data. Publication norms are enforced by the impact of publication on reputation. Reputation is central to the institution of open science. It provides direct rewards, which can be “cashed out” in terms of positions at preferred institutions and other recognitions of status, such as awards. Perhaps more importantly, the reputation system largely governs the allocation of funding resources. Reputation influences whether a researcher is successful in obtaining government funding via a mechanism administered largely by peer review. It also determines the extent to which a given researcher can influence both publication and funding decisions because reputation is a gateway to participation in the peer review process and in other forums that set research priorities at a more systemic level. Thus, though funding for academic research comes largely from government sources, the academic reputation system largely governs the way in which funding is deployed “on the ground” and influences decisions about funding allocation at more system levels.

Open science is a highly collaborative enterprise with relatively institutions. There are various plausible explanations for these institutional features, including the complexity of the problems themselves, the need for diverse expertise and perspectives in tackling scientific problems, and the fact that nature must be taken as it is and cannot be designed for modularity. Collaboration may also be necessary as result of the need to share expensive equipment. While there are undoubtedly some tasks in scientific research that can be allocated to individuals in small “chunks,” for the most part scientific research cannot rely on the crowdsourcing of low-cost tasks from dispersed individuals. It generally requires expertise that is specialized and costly to acquire. Scientific expertise also tends to be “asset specific” in that there is a high cost in moving from one research area to another. For these reasons, collaborations tend to be relatively longer term and are dependent on synergies between collaborators. Even outside of collaborations, relationships between scientists in the same field tend to involve repeated interactions. As a result of all of these features of open science, a reputation mechanism for allocating rewards is both feasible and sustainable.

Over the last few decades, and with the encouragement of the Bayh-Dole Act, scientific researchers have become increasingly likely to seek patents. Patenting by nonprofit researchers has been controversial and, in fact, community norms would frown on any attempt to enforce a patent against a fellow researcher. As a result, patents are employed mainly in “tech transfer,” the process of converting a scientific invention into a commercial product or service.

ii. User Innovator Communities

User innovators are incentivized by the direct benefits of use. They usually also have cost advantages because they have obtained information about desirable innovations as a side effect of use (and hence at no incremental cost). They are thus more likely than others to find it cost-effective to develop certain innovations. A user innovator community is one in which innovation is motivated by use and pooled in a system based on sharing norms and reputation-based rewards. In some cases, user innovators do not compete with one another and even enjoy the process of sharing information, meaning that the difficulties with free riding and secrecy that patents are intended to overcome simply do not arise. In others, user innovators compete with one another in some respects, but rely on norms rather than the patent system to allocate rewards

(usually reputational) and encourage disclosure. These communities' reasons for eschewing patents have not been studied thoroughly as an empirical matter, but the transaction costs associated with acquiring, enforcing and licensing patents, along with trust in the community's allocation of reputational rewards are likely explanations. Disseminating user innovations to non-innovating users often requires investments in standardization, refinement, manufacturing facilities, and so forth that user innovators have little incentive to undertake. Commercial firms (sometimes headed by user entrepreneurs) often take over these tasks and may seek patents on the results of those additional investments.

iii. Medical Procedure Innovation

Physicians have long maintained a system of sharing medical and surgical procedure inventions that could have been patented and licensed instead. The sharing regime is enforced by a system of reputational rewards, under-girded by a norm favoring publication and disclosure, and by formal ethical restrictions on exclusivity. The robustness of the sharing regime was demonstrated rather dramatically in the 1990s, when the medical community rose up against an attempt to enforce a patent on a particular surgical technique. The community lobbied for the abolition of medical procedure patents and eventually won a more limited exemption from infringement remedies.

I have argued in detail elsewhere that physician opposition to medical procedure patenting is best understood in light of the fact that medical practitioners are "user innovators" of procedures. Medical practitioners have intrinsic motivations to share their innovations: they generally are altruistically concerned with patient health. Sharing procedure innovations also allows for vetting, critique and improvement. Medical practitioners also compete with one another for patients, high-status positions, lecturing opportunities and so forth. Moreover, sharing procedure innovations is costly (often more costly than developing them), involving taking time to write articles, teach other physicians and so forth. Competition, combined with the costs of sharing, creates incentives to free ride on the common pool of knowledge. Rather than rely on patents, the medical procedure innovation system awards reputational rewards to those who share their innovations and imposes ethical sanctions against those who do not contribute their innovations to the common pool.

Though the physician procedure innovation system looks much like the open science research system in its reliance on publication, on reputation, and on social norms of sharing, there are important differences. Most obviously, procedural innovations are in general much less dependent than the open science research system on government funding. Often, the extra time and effort invested in procedural innovation is modest. Procedure innovations benefit physicians and their patients directly. A reputation for innovation may lead to greater market demand for a physician's skills.

B. Patents or an Alternative?

With these examples in mind, we can try to understand the circumstances under which an alternative innovation system is likely to be socially preferable to the patent system. The patent system has three main types of social costs: i) deadweight losses due to exclusivity, ii)

misalignment with socially valuable innovation directions because of insufficient or excessive demand resulting from myopia, inability to pay or externalities, and iii) transaction costs of awarding patents, defining boundaries, licensing and litigation. By comparing the patent system to alternative innovation systems, we can begin to map out the circumstances under which an alternative innovation system might be socially preferable. The comparison depends on both the technology (which determines, for example, how effective patents are as property) and the characteristics of the potential innovators (which determines the effectiveness and costs of alternative governance regimes). This section briefly describes five factors that are relevant to the comparison.

1. Intrinsic and user motivations for invention

As already mentioned, if first mover advantages and intrinsic or use motivations provide sufficient return on inventive investments, patents are socially wasteful. Nonetheless, individual inventors optimizing private returns may still obtain them. While secrecy is a potential concern whenever patents are not available, it is not an issue for all inventions. Some inventions are self-disclosing, while patents are not always needed to reward disclosure of inventions that result from intrinsic or use motivations. Disclosure may have its own rewards in terms of reputation, access to others' inventions, access to feedback, opportunities for collaboration, and intrinsic benefits of social and intellectual interaction. (Harhoff et al. 2003, see also Strandburg 2008).

2. High patent system transaction costs

Patents may also be socially wasteful in areas where they tend to impose particularly high transaction costs. The costs of defining claim boundaries and providing notice are especially high in some technological contexts. One also might expect high transaction costs when a large number of patent claims read on one commercial embodiment or when patent rights are awarded for inventions that have multiple uses and are difficult to design around. Transaction costs also are high wherever the scope of patent claims tends to be incommensurable with useful embodiments of the rights, resulting in overlap and recurring needs for renegotiation. If the patent system's transaction costs are high enough, they may even outweigh the value of the boost that patents provide to innovation in the field.

3. Assigning rewards for innovation

As is illustrated by the examples above, many groups of inventors allocate rewards and credit for means other than the patent system. Why might this be? First, while markets must rely on financial rewards, alternative innovation governance regimes often can confer reputation-based rewards that do not depend on exclusivity, thus avoiding its costs. Second, an alternative system may have lower transaction costs in some contexts, often by using a reward system that does not depend on putting precise boundaries around each individual's contributions. It may do a better job of allocating rewards by employing local and specialized knowledge. Reputational rewards are valuable only within the group that assigns them, however. Most inventors will require some means to cash out those rewards eventually through a better salary, consulting fees or the like. The feasibility and social desirability of a reputations-based reward system depends on the context in which a particular sort of invention is made.

4. Governance Mechanisms

To be successful, an alternative innovation system usually will need some workable means of evaluating contributions, distributing rewards, enforcing norms, and so forth. Depending on the context, there are a variety of governance approaches that may be successful. Some groups, such as high-end French chefs (Fauchart and von Hippel 2008) or extreme sports enthusiasts (Franke and Shah 2003), are small and cohesive enough that informal norms alone are effective. Other groups, such as Debian developers (Coleman 2012) or academic scientists (Strandburg 2009, Strandburg 2010 and references therein), have more structured governance regimes involving mechanisms for screening newcomers and inculcating the group’s norms, as well as enforcement mechanisms based on the group’s control of reputation and various other benefits of group membership. Still others, such as Wikipedia, adopt formalized governance and dispute resolution regimes (Mehra and Hoffman 2009). Private governance may not be effective for all technological areas, however.

5. Anti-competitive and anti-social effects

Privately ordered governance arrangements can also be socially costly to the extent they are designed to serve the vested (and potentially anti-competitive) interests of a group of insiders at the expense of the larger society. Alternative innovation arrangements may have the same pitfalls. As is illustrated by antitrust’s struggles with patent pools, it is not always easy to distinguish socially valuable innovation arrangements from socially detrimental cartels. Indeed, many institutional arrangements will have aspects of both. Government-administered patent systems can break up undesirable insider arrangements by provide a mechanism for outsiders to force their way into a creative arena. Government innovation arrangements, including the patent system and various forms of subsidies, are not immune from private interest distortions, of course, as is well known from the public choice discourse.

This list of factors is preliminary and undoubtedly incomplete. The tradeoffs associated with these and other factors will depend on the technological context. A complete and accurate cost-benefit comparison between the patent system and alternative innovation systems in any given context will never be easy and may not even be possible. Nonetheless, factors such as these can be used to guide a rough assessment of whether a patentable subject matter exclusion would be socially beneficial for some category of inventions.

C. Can We Avoid the Hard Choices?

Even if an alternative is preferable to the patent system in a given context, a patentable subject matter exclusion might not be necessary if the alternative system can coexist with the patent system. There are several circumstances under which a patent system might conflict with an alternative innovation arrangement. First, the availability of two innovation systems might allow inventors to “double dip,” imposing unnecessary social costs. If, for example, an inventor receives both upfront payment via a prize, procurement, or crowdsourcing, and a patent, the inventor will normally be overcompensated for her inventive investment.

Second, the availability of patents might disrupt the viability of an alternative system. For example, the availability of patents may decrease the costs of defecting from an alternative system, even causing the alternative system to unravel. This is particularly likely if the patent system's assignment of rewards is out of sync with the alternative system's assessment of the value of a given contribution. Those who have outside sources of rewards are less bound to the norms and rewards of the alternative institution. Patentholders may also be able to impose costs directly onto participants in an alternative institution by enforcing patents against those who have opted out of obtaining them for their own inventions. Participants in the alternative institution must either defend lawsuits, take licenses, or resort to defensive patenting. (One can see these dynamics playing out in the open source software arena.) In some cases, private ordering, through contracts, licenses, norms or more formal governance, may be sufficient to preserve the benefits of the alternative regime even when patents are available. Even when that is the case, the need to deal with the availability of patents generally makes the alternative regime more socially costly than it needs to be. A patentability exclusion avoids those costs.

Worries about the destabilizing effects and costs of patent availability have led to resistance against the introduction of patenting into some fields of technology. For example, as alluded to above, physicians reacted strongly when a fellow physician attempted to enforce patent claims on a medical procedure, lobbying for and obtaining a statutory change that abolished patent remedies for infringement of such patents by physicians.

D. Assessing the Tradeoffs

Despite the many uncertainties about the relative benefits of alternative innovation systems, the above examples and analysis illustrate that such institutions are feasible and sustainable and that there are substantial reasons to anticipate that they will be socially preferable to the patent system in at least some, and perhaps many, arenas. The easy solution of allowing patents and alternative mechanisms to coexist is no panacea, since co-existence may either be costly or lead to the unraveling of a socially preferable alternative. There are no certainties here. Nonetheless, the patentable subject matter doctrine unavoidably selects among innovation regimes. Ignoring that fact does not do away with the risk. Acknowledging that there are alternatives to the patent system will not make patentable subject matter determinations easy. Satisfactory empirical evidence of the relative costs and benefits of alternatives to the patent system may be unavailable at the time when decisions need to be made. This is not an unfamiliar position for patent policymakers. Empirical evidence of the value of the patent system itself remains inconclusive, though recent evidence strongly suggests that its value varies by technology. Moreover, the doctrine faces similar uncertainty in balancing market costs and benefits in its nonobviousness and scope doctrines. As the Court's focus on a realistic appraisal of the market in *KSR* reminds us, it is unwise to respond to uncertainty with formulaic rigidity. The better approach is reasoned analysis of the likely consequences in the real world of particular doctrinal choices. Such an analysis should be based in part on available empirical information, but will inevitably be based largely on more informal assessments and normative judgments.

One way to deal with the difficulty in assessing costs and benefits might be to adopt some sensible defaults. We know that the patent system unavoidably imposes deadweight exclusivity costs that raise consumer prices and tax downstream innovation. We also know that the

availability of patents sometimes threatens alternative innovation approaches. It might make sense to adopt a default position against awarding patents where viable alternatives are likely to be available. Where intrinsic or use motivations or first mover advantages are strong, for example, a patentable subject matter exclusion should be seriously considered, especially if a sensible category of inventions to exclude can be defined without too much cost. The debate about business method patents, for example, is largely framed in these terms, with those favoring the exclusion arguing that business method innovation is sufficiently motivated by first mover advantages and those opposing it arguing that the business methods category is too difficult to define properly. Reputation-based systems are particularly attractive because they encourage not only the invention, but also the disclosure and dissemination of novel technologies to others in the field. Reputation-based systems also avoid costs associated with defining precise patent claims boundaries and can rely on the local and specialized knowledge of those in the field, rather than the more general knowledge of PTO examiners, to assign inventive credit in a particular technological and institutional context. Thus, it may be especially important to consider patentable subject matter exclusions for categories of inventions for which a reputation-based innovation system is feasible. I do not mean to suggest here that a viable alternative will always be preferable to the patent system. Alternative systems have their own transaction costs and potential for socially harmful behavior. The current strong default favoring the availability of patents in all technological arenas is unjustified, however, once one realizes that patents are not the only mechanism for solving the free rider and other market failures that justify the awarding of exclusive rights in inventive ideas.

V. Patentable Subject Matter Doctrine from First Principles: Some Tentative Suggestions

Alternative innovation institutions are the “elephant in the room” of patentable subject matter doctrine. Courts cling to the traditional excluded categories of abstract ideas, natural phenomena and products of nature, but the market paradigm provides an inadequate basis for interpreting and applying them. The doctrine’s failure to acknowledge that it affects the *selection of innovation institution* prevents it from taking a reasoned approach to the issue and thus is at least partially responsible for the doctrine’s incoherence and lack of clarity. By submerging the institutional question, the current doctrine also fails to acknowledge the role that alternative innovation systems, most notably open science, already play *sub rosa* in the case law. The current doctrine’s formulaic approach to the basic exclusions also renders it incapable of adapting, as above analysis suggests that it should, when social institutions for technological innovation evolve. This Part attempts a tentative outline of a first principles patentable subject matter doctrine. It takes a categorical approach because comparisons between innovation regimes are inherently categorical. Innovation occurs in social contexts in which the patentability of one invention affects the innovation environment for related inventions. Tradeoffs between innovation regimes cannot be assessed on a claim by claim basis. An additional caveat is required: This article does not attempt to address the practical question of how best to implement the approach to patentable subject matter doctrine outlined here. While the courts may be capable of assessing the tradeoffs in many cases through the adversarial system, an expert agency might be better suited to the kind of policy determination required. These questions are extremely important, but are beyond the scope of this discussion, which focuses on the substance of the doctrine.

A. Defining Excluded Categories

Though the task is not straightforward, the goal in defining exclusions from patentable subject matter is now clear: to identify classes of inventions for which alternative innovation institutions are preferable to the patent system. This goal substantially shifts the analysis. Rather than begin by asking how to define “abstract idea,” “natural phenomenon,” or “product of nature,” the analysis begins with a much more basic assessment of how innovation proceeds (or can best proceed) in real world contexts. An obvious place to start is by studying both past and current alternative innovation institutions, both because they might suggest contexts in which patenting is unnecessary and because they provide clues as to when we might expect alternative approaches to be socially beneficial. We can also look for clues in the footprints of institutional concerns in the case law. We should also listen to the voices of inventors in a particular technology space, particularly when they oppose patents or raise concerns about their effects. And of course we should maintain a critical perspective with respect to any potentially excludable category suggested by these inquiries because of the potential costs already discussed. Finally, when we view patentable subject matter doctrine as grounded in real world social arrangements, we should be open to the possibility that the excluded categories will need to evolve over time. This is an uncomfortable recognition for those of us steeped in the essentialist approach of current doctrine, but it is not a terribly surprising feature of a system designed with innovation in mind.

This is certainly not the place for an exhaustive inquiry into potentially excludable categories, but a few examples will illustrate the type of analysis I have in mind.

1. Business Methods

Many scholars, policymakers, and, in *Bilski*, four Supreme Court Justices, have argued that business methods should be unpatentable, in large part, because patents are unnecessary to promote business method innovation in a competitive market and thus impose unnecessary costs. The argument, in essence, is that for business methods the traditional competitive market is superior. The approach to patentable subject matter outlined here would bring this debate into the foreground, rather than miring the discussion in abstractions about abstract ideas.

2. User Innovator Communities

The existence of a strong user innovator community is an important clue that a patentable subject matter exclusion may be appropriate for the relevant technology. The most important reason for this is that user innovator communities largely internalize both the costs and the benefits of the innovation institution in which they function. While they benefit, as inventors, from private returns to patent exclusivity, they also bear the deadweight losses and transaction costs of patents as downstream inventors and, most importantly, users, of others’ inventions within their field. From this perspective, it is not surprising that many of our most successful and longstanding alternative innovation systems involve significant user innovation. Examples include scientific researchers (as user innovators of both research tools and scientific knowledge itself), medical professionals (as user innovators of medical and diagnostic procedures), and software

programmers (where the most successful open source projects are infrastructural programs used by programmers), tax professionals (who are the primary users and inventors of automated tax strategies).

User innovator communities tend to resist patenting. Indeed, each of the groups I mentioned has (with greater or lesser success) resisted the extension of patenting into its bailiwick on the grounds that patents would disrupt a working alternative system. Resistance to patenting by user innovators should get our attention. Indeed, if a user innovator community were a completely closed system, the community’s opposition to patents should be decisive. Most cases are more complicated because user innovation often has externalities. In each of the cases described above, for example, the externalities are fairly significant. The degree to which a user innovator community’s interests are aligned with those of the larger society will vary and should temper the persuasiveness of the community’s position on patents. Nonetheless, we would be foolish to ignore the views of such a community without good reason.

3. Software Patents

Software patents have been controversial throughout their history and are particularly under debate today. At least according to some commentators, we now are dealing with an extremely expensive software patent arms race, costly litigation, and a “patent troll” problem, all of which could have been avoided by a patentable subject matter exclusion. A more detailed assessment would be needed to draw a conclusion about whether institutional analysis would support a ban on software patents (or how one would best define an excluded category). It is clear, however, that such an analysis would differ dramatically from the rather inscrutable analysis of whether software claims are “abstract ideas.” Such an analysis would consider, for example, the historical opposition by the programming community to the extension of patents to software in light of the extent to which that community is engaged in user innovation, the high transaction costs imposed by the fact that software tends to be both complex and cumulative, the importance of first mover advantages and intrinsic motivations in software innovation, and the availability of alternative innovation systems, such as copyright, trade secrecy, and open source. Moreover, a pragmatic analysis of this sort might suggest useful distinctions, such as between the infrastructural programs for which user innovation predominates and consumer-facing software, that could inform the definition of a patent ineligible category.

4. Natural Phenomena

Many of the Supreme Court’s patentable subject matter cases involve the traditional exclusion for natural phenomena. The rationale for the natural phenomenon exclusion is unclear. As discussed above, the Court sometimes relies on a broad downstream impact rationale, but the awkward fit of that justification to a category that includes many relatively narrow phenomena is evident from the Court’s discussion of the issue in *Mayo*. Discussions based on natural phenomena as basic tools of innovation are more persuasive, given transaction costs concerns, but like the downstream impact thread generally, fail to explain how these important tools will be invented without patent incentives. Other discussion in the case law rely on metaphysical references to natural phenomena as part of a “storehouse of knowledge,” often assumed to pre-

exist human intervention, that should be available to all. While such arguments may have moral force, they have little purchase in a utilitarian analysis.

The institutional approach provides a far more compelling justification for the natural phenomena exclusion. Over time, society has developed – and subsidized – the “open science” system for incentivizing the discovery, disclosure, and dissemination of observations of natural phenomena, often denoted “open science.” The system overcomes the free rider problem by relying on publication as a prerequisite for reputation-based rewards and as a means of deterring secrecy. The open science system allocates public research funding using a complicated amalgam of peer review, agency expertise, legislation, private philanthropy, and corporate donations. It is governed partly by informal social norms and partly by rules and other governance mechanisms imposed by funders and employers.

Transaction costs associated with patents on basic natural phenomena are likely to be huge for two distinct reasons. Some “laws of nature” are broad general principles with diverse applications, patents on which are likely to have high transaction costs for the reasons discussed in association with the downstream impact thread of patentable subject matter doctrine. The transaction costs associated with patents on narrower natural phenomena are also likely to be high, however, at least in the aggregate. Basic natural phenomena are highly interrelated; answering a given practical question frequently requires weaving a number of scientific observations together. Natural phenomena also tend to be cumulative on a very fine-grained scale. Science for the most part advances by small steps. Finally, scientific observations of natural phenomena are simultaneously hard to design around and tentative. It is generally very difficult to falsify or demonstrate the limitations of a particular scientific observation without employing that observation in conjunction with others. Patents covering natural phenomena thus have the potential to block progress because the patentee might have little incentive to test the accuracy of the observation, while others would be prohibited from doing so.

In sum, because of the upstream, interrelated and cumulative nature of scientific discovery, the patent system is likely a costly means to incentivize the discovery, disclosure and dissemination of observations of natural phenomena in comparison to the reputation-based open science system. As discussed earlier, the open science system also directs research effort in directions that would not be adequately induced by market demand because of long time horizon and/or uncertainty of practical payoff, widely dispersed and difficult to aggregate benefits (collective action problems), or social goals that are not reflected in market willingness to pay.

At times, the Supreme Court has recognized the relationship of the natural phenomena exclusion to the institution of open science. In *O'Reilly v. Morse*, for example, the Court noted that “for some years before Morse made his invention, scientific men in different parts of Europe were earnestly engaged in the same pursuit. Electro-magnetism itself was a recent discovery, and opened to them a new and unexplored field for their labors, and minds of a high order were engaged in developing its power and the purposes to which it might be applied.” Moreover, “very soon after the discovery … it was believed by men of science that this newly-discovered power might be used to communicate intelligence to distant places.” The Court also considered and rejected an argument that “the inquiries [Morse] made, or the information or advice he received, from men of science in the course of his researches” should disqualify him from

receiving a patent. In fact, Morse’s patent was obtained during a widespread effort of the scientific community to develop telegraphy and generated a dispute about priority of invention.

While the Court’s rejection of Morse’s broadest claim certainly was based on its breadth, the Court’s concern with the downstream impact of the claim was closely tied the claim’s potential impact on the scientific community:

[W]hile he shuts the door against inventions of other persons, the patentee would be able to avail himself of new discoveries in the properties and powers of electromagnetism which scientific men might bring to light. … New discoveries in physical science may enable him to combine it with new agents and new elements, and by that means attain the object in a manner superior to the present process and altogether different from it. And if he can secure the exclusive use by his present patent he may vary it with every new discovery and development of the science, and need place no description of the new manner, process, or machinery, upon the records of the patent office. And when his patent expires, the public must apply to him to learn what it is.

In other words, Morse’s broadest claim was problematic in part because it would allow him to take too much advantage of freely available scientific developments without reciprocal disclosure.

Mid-20th Century cases begin to justify the natural phenomena exclusion on the basis that natural phenomena are “basic tools of scientific and technological work,” but for the most part ignore the institutions in which scientific work takes place, continuing to refer to scientific principles and even algorithms as “mere’ recognition of a theretofore existing phenomenon or relationship,” providing no explanation of why the often costly effort involved in scientific research need not be motivated by a patent payoff. Very recently, however, as discussed in the next Part, the Court in *Mayo* took important steps toward acknowledge the role the institution of open science plays as a back-up for the natural phenomena exclusion.

The institutional rationale provides a far sounder basis the natural phenomena exclusion than essentialist rationales based on defining “nature.” It provides a rationale for determining the contours of the “natural phenomena” exclusion that explains how discoveries of natural phenomena will be incentivized. It also provides a basis for the exclusion to evolve, if necessary, to accommodate changes in the subject matter to which the open science system applies.

B. Step Two: A First Principles Perspective on Determining Whether A Claim Incorporating Categorically Excluded Subject Matter is Patentable

The second stage rule for determining whether a claim incorporating categorically excluded material is patentable should be crafted in light of the justification for the categorical exclusion. This straightforward requirement, in and of itself, would go far to rationalize patentable subject matter doctrine, whether one adopts an institutional approach or not. Unfortunately, while courts have devoted considerable attention to the second stage question, they have not for the most part focused on matching the second stage rule to the rationale for the exclusion at issue. This is perhaps inevitable given that the theoretical understanding of the traditional exclusions is in such

a muddle, but it exacerbates the problems with the doctrine. The analysis has been particularly confused by attempts to insert consideration of “preemption” into the second step. Whether a claim has overly broad downstream impact can be assessed in a single step by looking at the claim as a whole. An inquiry into “preemption” at the second step of a categorical exclusion analysis makes no sense.

The analysis in *Bilski* illustrates the problem. In *Bilski*, the Court first concluded that “risk hedging” is an unpatentable abstract idea because patenting such a “fundamental economic practice” would “preempt use of this approach in all fields.” It then considered narrower claims limited to the use of hedging in “the energy market,” and requiring that certain “random analysis techniques” be used to “establish some of the inputs into the equation.” Here the analysis (such as it was) goes off the rails. If risk hedging is unpatentable because of its broad downstream impact, surely the relevant question should be whether the narrower claims still have broad downstream impact despite these limitations. But the Court does not ask that straightforward question. Instead, the Court opines that the claims are unpatentable because the additional elements merely “limit[] an abstract idea to one field of use or add[] token postsolution components.” The Court never asks whether the claim still has broad downstream impact despite requiring that particular techniques be used to determine inputs. Suppose, as seems plausible and is suggested by the dissent, that those limitations on how the inputs are determined mean that the downstream impact of the claim is no longer broad. What then is the justification for the unpatentability determination? The Court’s analysis provides no clue. The problem here is precisely that there is no apparent logical connection between the broad downstream impact rationale for excluding “risk hedging” and the second stage rule that “token” postsolution components cannot impart patentability to the claim.

From an institutional perspective, the relevant question at the second step is whether the patent system is needed to induce the claimed invention, given the work that will be performed by the alternative innovation institution on which the categorical exclusion is based. For example, consider a claim to a part for a wind turbine that incorporates a high temperature superconducting material. The “natural phenomena” in this claim are the superconducting properties of the material. The “step two” question, then, is whether we can expect the wind turbine part to be invented as part of “open science” or whether a patent is required. In this case, it is evident that a wind turbine part is a commercial product that we do not expect to emerge from “open science.” Thus, the claim is addressed to patentable subject matter. Now consider a different claim, to a magnet used for research purposes that uses the newly discovered superconducting material. The institutional question here is more difficult. We would need to know a bit more about whether scientists invent such magnets as research tools or they typically are designed by engineers working in commercial firms. Finally, consider a third example, in which a lump of the superconducting material is packaged with tongs and a small liquid nitrogen Dewar to be used in high school physics labs. Assume also that the material was newly discovered by the scientist who filed the patent claim. This kit is unlikely to be a normal output of the open science system, but we might nonetheless decide that this claim is unpatentable because the output of the open science system enables educational suppliers or high school teachers themselves to make such a kit without any inventive effort. Note that, because the material was newly discovered by the scientist, even this claim would not have been deemed

obvious because obviousness doctrine assumes that all incentives – including, in this case, the incentives to discover the material – must be provided by the market.

As the previous examples illustrate, from the institutional perspective, the second step of the patentable subject matter doctrine is essentially a question about “tech transfer” from one innovation institution (governing the excluded elements) to another (the patent system). A claim should be patentable if the patent system is necessary to induce this “tech transfer.”

C. *Mayo v. Prometheus* through an Institutional Lens

While the Supreme Court has not explicitly adopted an institutional approach to the patentable subject matter question, it comes closest to following such an approach implicitly in its opinion in *Mayo v. Prometheus*. This section thus uses the *Mayo* case as a foil for illustrating in more detail how an institutional approach to patent would work out in practice. In *Mayo*, the Supreme Court surprised most observers by holding unanimously that the natural phenomenon exclusion from patentable subject matter invalidated claims exemplified by the following:

“A method of optimizing therapeutic efficacy for treatment of an immune-mediated gastrointestinal disorder, comprising:
(a) administering a drug providing [a particular metabolite] to a subject having said immune-mediated gastrointestinal disorder; and
(b) determining the level of [the metabolite] in said subject having said immune-mediated gastrointestinal disorder,
Wherein the level of [the metabolite] less than about [a particular concentration in the blood] indicates a need to increase the amount of said drug subsequently administered to said subject and
Wherein the level of [the metabolite] greater than about [a particular concentration in the blood] indicates a need to decrease the amount of said drug subsequently administered to said subject.”

The claims were based on the results of a research study correlating patients’ measured levels of a particular drug metabolite with their responses to the drug. Using this statistical correlation, the patentees devised guidelines to inform physicians when they should be concerned that the dosage was too high or too low. To infringe the claims, physicians had only to consider adjusting dosage in light of the guidelines. In the litigation, Prometheus sued the Mayo Clinic for inducing physician infringement by performing laboratory analysis of the metabolite level and instructing doctors as to the guidelines.

The Court subjected the claims to a two-step analysis: First, the Court held that the correlations were unpatentable laws of nature. Second, the court held that the additional steps in the process did not render the claims patentable because they “consist[ed] of well-understood, routine, conventional activity already engaged in by the scientific community” and thus added no “inventive concept” over and above the discovery of the natural correlations.

1. Defining and Justifying the Excluded Natural Phenomena Category

In determining that the claimed correlations were “natural phenomena,” the Court rejected an argument that correlations between drug toxicity or efficacy and the blood level of a chemical produced by the metabolism of the drug were not “natural phenomena” because no metabolite would have been present without the manufacture of the drug itself and the human activity of ingesting it. In a passage fairly typical for natural phenomena opinions, the Court differentiated between the human action that “trigger[s] a manifestation of [the correlation] in a particular person” and “the relation itself,” which “exists in principle apart from any human action.” Court “emphasized … a concern that patent law not inhibit further discovery by improperly tying up the future use of laws of nature,” noting that such laws are “the basic tools of scientific and technological work.”

Elsewhere in the opinion, however, the opinion took on a more institutional flavor. It broke new ground by its forthright acknowledgement that the natural phenomena category extends to both narrow and broad laws of nature, consistent with its characterization of the issue as a “building block” concern. Though few natural phenomena are broadly foundational, in the sense of Newton’s laws, essentially all building blocks, in that they are combined and recombined in the incremental and cumulative advance of scientific understanding. Because natural phenomena are so intertwined, even narrow patent claims could result in a “vast thicket of exclusive rights over the use of critical scientific data … .” Thus *Mayo* described natural phenomena as being ill-suited for the patent system because of their overlapping and ill-defined boundaries, the need to aggregate many discoveries in order to make progress, and the lack of options for designing around them. These are the kinds of characteristics that provide a rationale for the open science regime.

Moreover, though it did not explicitly adopt an institutional perspective, the *Mayo* opinion clearly assumes the existence of an open medical/scientific research community that will discover correlations of the sort reflected in the claims. The Court noted that the claims at issue might “inhibit the development of more refined treatment recommendations … that combine [the claimed] correlations with later discovered features of metabolites, human physiology or individual patient characteristics.” Patents on natural phenomena are problematic because they would “require[e] potential users to conduct costly and time-consuming searches of existing patents and pending patent applications, and … the negotiation of complex licensing arrangements.” Moreover, the Court pointed out that the claimed discoveries were part of an ongoing thread of scientific research:

At the time the discoveries embodied in the patents were made, scientists already understood that the levels in a patient’s blood of certain metabolites … were correlated with the likelihood that a particular dosage … could cause harm or prove ineffective. … Indeed, scientists routinely measured metabolites as part of their investigations into the relationships between metabolite levels and efficacy and toxicity of thiopurine compounds.

An institutional approach to patentable subject matter analysis would compare the open medical/scientific research system to the patent system more explicitly, but the analysis would be informed by many of the same considerations that drove the Court to hold that the correlations at issue in *Mayo* are unpatentable natural phenomena. An institutional analysis also would have

avoided the unconvincing detours that the opinion took into attempting to justify excluding narrow natural phenomena on a preemption-type basis. In this particular case it would not even have been necessary to speculate as to whether the open science approach *could* have induced the discovery of the correlations at issue, since the patentees were academic researchers, who published the results of their clinical study in a medical journal. The advantage of an explicitly institutional approach in this case thus would have been to clarify and sharpen the analysis, avoiding detours into metaphysical debate about what is natural or into irrelevant preemption rhetoric.

2. The Second Step Analysis and *Mayo*'s Inventive Concept Rule

Having determined that the claims at issue in *Mayo* incorporate unpatentable natural phenomena, the next step is to determine whether they apply those phenomena in a patentable way. At this second step, the *Mayo* opinion requires that the claims demonstrate an “inventive concept” over and above the discovery of the natural phenomena. The claims at issue were deemed unpatentable because “apart from the natural laws themselves” they “involve well-understood, routine, conventional activity previously engaged in by researchers in the field.”

Applying an institutional approach, the second step should be aimed at determining whether the patent system was needed to induce the claimed applications of the natural correlations. The *Mayo* second step is quite sensible from this perspective and the end result is almost trivially correct. The medical/scientific research system is effective at incentivizing the discovery of natural phenomena, as well as their disclosure and dissemination *to other members of the medical/scientific community and readers of the scientific literature*. In *Mayo*, the claimed application of the natural phenomena for optimizing drug dosage was well within the skill set of the physicians who are the audience for the scientific literature in which the phenomena would be reported. Because the application was routine was for those who would practice it, there was no need for a patent incentive to “transfer” the discovery to a commercial market.

The Court’s discussion of the second step makes much the same point: “[The] claims inform a relevant audience [doctors who treat patients with certain diseases with thiopurine drugs] about certain laws of nature; any additional steps consist of well-understood, routine, conventional activity already engaged in by the scientific community; and those steps, when viewed as a whole, add nothing significant beyond the sum of their parts taken separately. For these reasons we believe that the steps are not sufficient to transform unpatentable natural correlations into patentable applications of those regularities.”

Elsewhere, the Court explains further:

[T]he “administering” step simply refers to the relevant audience, namely doctors who treat patients with certain diseases with thiopurine drugs. That audience is a pre-existing audience; doctors used thiopurine drugs to treat patients suffering from autoimmune disorders long before anyone asserted these claims. . . . Second, the “wherein” clauses simply . . . tell the relevant audience about the laws while trusting them to use those laws appropriately where they are relevant to their decision-

making (rather like Einstein telling linear accelerator operators about his basic law and then trusting them to use it where relevant).

...

Beyond picking out the relevant audience, ... the claim simply tells doctors to: (1) measure (somehow) the current level of the relevant metabolite, (2) use particular (unpatentable) laws of nature (which the claim sets forth) to calculate the current toxicity/inefficacy limits, and (3) reconsider the drug dosage in light of the law. These instructions add nothing specific to the laws of nature other than what is well-understood, routine, conventional activity, previously engaged in by those in the field. And since they are steps that must be taken in order to apply the laws in question, the effect is simply to tell doctors to apply the law somehow when treating their patients.

In sum, the *Mayo* analysis is consistent with an institutional approach to patentable subject matter determination. Its inventive concept test determined that the patent system was unnecessary to induce the claimed inventions because the scientific discoveries on which they were based were expected output of the medical/scientific research system and were applied in a way that would be routine for the ordinary consumers of that output.

VII. Conclusion

Because patentable subject matter debates have generally ignored the institutional choices implicit in the doctrine, they tend either to focus on arcane philosophical questions about the meaning of “nature” or to become scholastically obsessed with the language of precedent, rather than its logic. The analysis I have sketched out here takes a first principles approach to the patentable subject matter question, focusing both on the downstream transaction costs that concern courts applying the doctrine and on the need to incentivize upstream innovation that concerns many of the doctrine’s critics. By recognizing that alternative innovation regimes may in some circumstances be socially preferable the patent system, the approach outlined here avoids metaphysical tangles and highlights pragmatic choices.

In the case of the natural phenomena exclusion this approach would incorporate many of the concerns that are scattered throughout judicial opinions applying current doctrine, but would provide a much more rational analysis and eliminate the confusing detours that currently muddy the analysis. Though the shadow of the open science system lies heavily over case law applying the natural phenomena exclusion, the opinions are confusing because they do not explicitly seek to determine whether the purported natural phenomenon is a likely output of the open science system and whether the patent system is needed to induce the claimed application of the phenomenon.

Though this Article takes the natural phenomena exclusion as its primary example, its general approach is broadly applicable. Indeed there may be many arenas, some of which are suggested here, in which alternative innovation regimes are socially preferable to the patent system. Patentable subject matter exclusions may be appropriate for those categories of inventions. The

identification and analysis of candidates for patentable subject matter exclusion, including the traditional abstract ideas and product of nature categories, is a subject for further research.

The institutional approach also suggests that patentable subject matter exclusions should not be viewed from a timeless, essentialist perspective, but rather from a perspective that takes account of the changing realities of the innovation environment. This suggestion is not as radical as it sounds. Because innovation institutions change over time, it should not be terribly surprising if patentable subject matter exclusions also evolve over time, sometimes expanding and sometimes contracting.