Applied Economic Perspectives and Policy (2013) volume 35, number 2, pp. 206–241.
doi:10.1093/aepp/ppt012
Advance Access publication on 25 March 2013.

Featured Article

Incentives for Innovation: Patents, Prizes, and Research Contracts

Matthew S. Clancy and GianCarlo Moschini*

Matthew S. Clancy and GianCarlo Moschini are with the Department of Economics and Center for Agricultural and Rural Development, Iowa State University.

*Correspondence may be sent to: moschini@iastate.edu.

Submitted 22 November 2012; accepted 9 February 2013.

Abstract *Innovation is essential for sustaining growth and economic development in a world that faces population increase, natural resource depletion, and environmental challenges. Incentives play a critical role in innovation because the required research and development activities are costly, and the resulting knowledge has the attributes of a public good. This paper discusses the economics of institutions and policies meant to provide incentives for research and innovation, and focuses on intellectual property rights, specifically patents, contracted research (for example grants), and innovation prizes. The main economic implications of these institutions are discussed, with particular attention paid to open questions and recent contributions.*

Key words: Competitive grants, Growth, Incentives, Innovation prizes, Intellectual property, Monopoly, Patents, Public goods, Research contracts, Research and development, Science policy, Tournaments.

JEL codes: O31, O34, O38.

Introduction

Research is a unique activity—it seeks to produce new knowledge. Such an intangible output is typically not of interest to final consumers per se, but it is best thought of as a (critical) intermediate input in the production of innovations: new products and new processes/technologies of direct interest to firms and consumers. Innovations constitute a distinctive attribute of modern economies. Indeed, they are deemed essential for the progress and growth that has been sustained in developed economies since the dawn of the industrial revolution. Much has been written about the process of research and innovation, elucidating the function of science, the role of public institutions and private concerns, and the critical parts played by a number of actors ranging from scientists and inventors who produce basic breakthroughs, to entrepreneurs who bring innovations to market. What emerges is a complex system fraught with market failures, with an acute need for public policies and coordinated actions which, however, are often beset by unintended consequences (Aghion, David, and Foray 2009). One thing is clear in this context, however: incentives matter.

Our focus on incentives is motivated by a quintessential characteristic of knowledge: it is intrinsically a public good. That is, it takes the form of information and as such it is nondepletable (nonrival in consumption), and absent specific institutional provisions, it is also nonexcludable. The free-rider externality that arises with public goods implies that its private provision is inefficient. Put differently, providers of this public good face an appropriability problem, one of the instances of market failures identified by Arrow (1962) in this context. Agents who have the possibility of producing new knowledge and innovations that are potentially of high value to society face the real prospect of being able to appropriate only a negligible portion of the value they produce. Insofar as it is costly to engage in inventive and creative activities, there is no presumption that a competitive allocation of resources in a market economy would lead to an efficient outcome. Specific concerns relate to the level and variety of this public good: left to itself, a market economy might devote insufficient resources to the production of new knowledge and innovations. Also, to the extent that heterogeneity exists regarding how nonappropriability relates to potential innovations, a suboptimal portfolio of innovations might be attained.

A number of solutions to the appropriability problem are possible, in principle. A standard prescription for the provision of a pure public good (e.g., national defense) is to assign production responsibilities to the government. In fact, governments in developed countries and emerging economies devote large resources to funding research and development (R&D) activities, with particular emphasis placed on basic research. Some of this government-funded R&D is carried out in government laboratories and research centers, but the bulk of it is outsourced to academic institutions and the private sector. The Appendix provides some data, mostly for the United States, that quantifies the scope of this enterprise. Such delegation of R&D performance takes on a number of forms, including competitive grants, procurement contracts, and the direct allocation of funds to performing units. Regardless of the performing units, ultimately the research work is the responsibility of scientists and engineers working either individually or as part of teams. Because research work relies on very specialized knowledge and skills that are heterogeneously distributed, and typically unobservable, and because the production of knowledge is inherently risky, asymmetric information distribution and moral hazard are common in this setting. The question of how to structure a set of incentives that provide effective motivation to research workers is critical. Accordingly, a major objective of this article is to discuss the incentive issues of common funding mechanisms for contracted research.

Government involvement in research and scientific activities leading to the production of new knowledge has a long history, but the extent of this participation has become particularly sizeable only during the last fifty years. In the United States, publicly funded R&D grew rapidly after World War II, but as a share of Gross Domestic Product (GDP) it has been declining since the mid-1960s (see the Appendix for further discussion). Going forward, it seems unrealistic to envision a much larger government role in this setting, particularly given the fiscal pressures faced by most developed economies. Thus, the private sector will most likely continue to contribute the lion's share of research funds. Indeed, as documented in the Appendix, industry-sponsored R&D in 2009 was twice the size of government-sponsored R&D. Such industry R&D, of course, benefits directly from government support through a variety of tax credits (in the United States, both federal and state credits are often available). Still, industry R&D investments must overcome the problems associated with the private provision of what could largely be a public good. A solution to the appropriability problem meant to work in a market setting takes the form of legally sanctioned intellectual property rights (IPRs), such as patents, copyrights, trademarks, and trade secrets. Another major objective of this paper, therefore, is to consider the nature of the incentive effects provided by IPRs. We focus specifically on patents, and investigate key ways through which they provide incentives for private research and innovation endeavors.

The (partial) privatization of the knowledge commons brought about by IPRs contributes to fostering the market provision of innovations, but what results is clearly a second-best solution. Indeed, some of the unintended consequences and side effects entail social costs that counter the presumably positive incentive effects that patents provide. Thus, there is a longstanding interest for alternative methods to provide basic innovation incentives to individuals and firms that are capable of leveraging their unique knowledge base and capabilities, but without incurring the deadweight welfare losses of patents. In recent years, an area that has shown a resurgence of interest in this context is that of innovation prizes. Thus, the last section of the paper addresses the main economic issues related to prizes, with some attention given to how they relate to the other incentive mechanisms and institutions analyzed earlier.

Patents

Patents go to the heart of the aforementioned appropriability problem by granting inventors the right to exclude others from the commercial exploitation (making, using, or selling) of the patented innovation. Thus, a patent endows the inventor with property rights on her discoveries, thereby affecting the excludability attributes of an otherwise pure public good. Such rights are limited in time (the patent's length) and scope (the specific claims granted by the patent). For an innovation to be patented, it must be *novel* relative to the prior art. It must also be *nonobvious* to a person with ordinary skills in the field (i.e., it must involve an inventive step), and it must be *useful* (i.e., allow the solution of an explicit problem in at least one application). To obtain a patent, its claims must successfully pass the examination of a governmental office, for example, the U.S. Patent and Trademarks Office (USPTO).[1] A host of other considerations apply (Scotchmer 2004, chapter 3). For example, what is deemed patentable subject matter has changed over time, and differs somewhat across

---

[1] *IPRs are the prerogative of national jurisdictions, although convention and treaties harmonizing such rights have a long history. A significant development in the international dimension of IPRs is the TRIPS agreement, which requires nations to meet minimum standards of IPR protection as part of their commitments for membership in the World Trade Organization (Moschini 2004).*

jurisdictions.[2] An important element of the patent institution is *disclosure*, which requires the patent application to describe the claims in sufficient detail to *enable* an individual skilled in the field to practice the invention.

The incentive potential of patents is due to their private value, which depends, *inter alia*, on the length of the patent grant and on the scope of the patent. The length of the patent is codified by law (twenty years from filing the application), although the effective economic life of the patent can be considerably shorter,[3] and influences how long competitors can be excluded from a particular market. The scope of a patent is more subtle and concerns the breadth of its applications, which relates to the range of products or processes that can be excluded by a patent's right (by the so-called doctrine of equivalents, a product might be found to infringe on a patented product even if it is not an exact replica). Unlike length, the breadth of a patent cannot be explicitly codified, and it is left to be determined by the patent's claims, as approved by the patent examiners and ultimately adjudicated by the courts.

The economic benefits and costs of the patent institution have been the object of extensive research, too much for us to review in a comprehensive and systematic manner. The reader is referred to existing compendia, of varying detail, which include Langinier and Moschini (2002), Scotchmer (2004), and Rockett (2010). Instead, in this section we discuss some of the patent institution's key features and implications, with an emphasis on issues that remain unsettled.

Patents' Basic Incentive to Innovate

Property confers privileges, and the pursuit of property has long motivated and energized individuals. The owner of a patent obtains a legally-sanctioned (potential) monopoly power,[4] just as the owner of any piece of property does. The prospect of being able to profit from such a monopoly can be a powerful incentive for would-be innovators, enough to spur efforts and investments that would not otherwise occur. To illustrate, consider the market for a new product that, once developed, can be produced at a constant marginal cost $c < \bar{p}$ (figure 1). The downward sloping demand $D(p)$ for this innovation, emanating from the choke-off price $\bar{p}$, displays the diminishing willingness to pay of consumers (arising, say, from their heterogeneity vis-à-vis income). A firm with a monopoly in this market, as endowed by a patent on the product, and absent opportunities for price discrimination, would charge price p^M, sell quantity q^M, and reap a per-period profit of π. Consumers enjoy a net surplus of S, and so

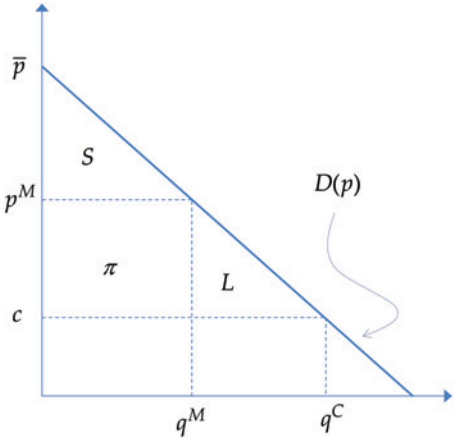
---

[2] *In the United States an expansive view of what can be patentable has evolved, as aptly captured by the oft-quoted remark included in the* Diamond v. Chakrabarty *1980 U.S. Supreme Court decision: "anything under the sun that is made by man." Still, discoveries about natural phenomena, mental processes, and abstract concepts are typically not patentable subject matter.*

[3] *For new pharmaceutical drugs, for example, the most recent estimate of the average length of patent exclusivity is 12.4 years (Grabowski et al. 2011). More generally, the period over which an innovation can command supra-competitive pricing must contend with the process of "creative destruction," whereby innovation leads one monopoly to supplant another (Schumpeter 1942).*

[4] *A patent does not actually grant the right to produce or sell something. For example, a patented pharmaceutical still needs to clear regulatory hurdles before being allowed on the market; a patented innovation might actually infringe on other patented products; or there might not be a market for the patented product or process, which might be novel but inferior to existing ones.*

Figure 1 Patents and Monopoly Pricing



both producer and consumers have immediate benefits from the patented product. Upon expiration of the patent, competitive production is presumed to eliminate profit and increase consumer surplus by the quantity $\pi + L$. The prospect of earning the profit flow π for the duration of the patent can be a powerful incentive to invest in R&D. If the present value of such a profit stream exceeds the expected R&D costs of developing the new product, a risk-neutral firm with a unique opportunity to pursue this project would find it desirable to invest.

What figure 1 also illustrates is a downside of the patent grant; *ex post*, that is, after the new product is developed, q^M is inefficiently low because it excludes consumers whose willingness to pay exceeds the marginal production cost. The fact that the provision of the good is below the efficient level q^C means that there is a per-period deadweight welfare loss of L. This illustrates the tradeoff between static inefficiency and dynamic efficiency engineered by patents (Nordhaus 1969). From an *ex ante* perspective, patents are a necessary incentive that may bring about innovations that would not otherwise take place. Insofar as innovations are underprovided in the counterfactual market equilibrium (i.e., without patents), the incentive to innovate provided by patents should ameliorate the market allocation of resources to innovative and creative activities. *Ex post*, however, the (limited) monopoly position established by patents entails an inefficiently low production of the new product, that is, first-best allocations are not achieved. Furthermore, patents may not provide enough incentive for the would-be innovator, because only a fraction of the realized benefits can be appropriated. For the example of figure 1, it may be that the innovation cost exceeds the expected profit of a patentee monopolist, but it is lower than the flow of total surplus. When individuals are endowed with unique innovation opportunities ("ideas" that are costly to develop, as in Scotchmer [2004]), such socially desirable innovations would not be pursued, that is, there remains an underprovision of innovations. On the other hand, when an innovation opportunity is open to many individuals or firms (e.g., a drug to treat a known condition), the prospect of winning the patent race that arises in this setting can lead to overinvestment in research (excessive entry into the R&D contest).[5]

As noted earlier, the extent of the innovation incentive provided by a patent depends on its duration (the patent's length) and on the scope of the exclusionary rights granted (the patent's breadth). The latter is critical because it directly affects both the size of the innovator's profit flow π, and the deadweight welfare loss L. Narrow patents (e.g., easy to imitate or to invent around) may provide insufficient incentives to innovators, but broad patents (e.g., providing exclusionary power over a large set of applications) may induce excessive deadweight welfare losses. The optimal mix of patent breadth and length, therefore, is of crucial importance for the patents' goal of balancing static and dynamic efficiency effects. Notwithstanding valuable insights, however, extant work does not provide results of general applicability. Gilbert and Shapiro (1990) find that narrow infinitely-lived patents are optimal, but Klemperer (1990) exhibits a model where either narrow infinitely-lived patents or broad patents of minimum length may be optimal, and Gallini (1992) finds support for short patents (Denicolo [1996] provides a unifying framework for these results to emerge).

More General Innovation Contexts

Beyond their impact on appropriability, patents have additional relevant economic effects that indirectly influence the incentive to innovate. First, the disclosure requirement of the patent grant helps the dissemination of knowledge. Disclosure furthers research efforts in directions that avoid duplicating existing results. Also, knowledge of the nature and feasibility of patented innovations may spur new research efforts that build on the full current state of knowledge. Such results are best appreciated when the alternative to patent protection is that of trade secrets (also sanctioned by law). Second, insofar as an inventor is not best positioned to commercially exploit an innovation, or when an innovation has utility in many applications, patents can facilitate technology transfer by allowing efficient contracting and licensing. The property right granted by a patent is essential in this setting to overcome the predicaments of developing a market for new knowledge *qua* information, noted by Arrow (1962). The ability to effectively license an innovation can raise its expected value and thus strengthen the *ex ante* incentive to innovate.[6]

The incentive effects of patents are more subtle when inventions are seen as part of a cumulative process whereby new products and processes are the springboard for more innovations and discoveries. Patents on early inventions may restrict access and discourage follow-on innovations. In such a setting it is possible that patents may actually slow innovation (Bessen and Maskin 2009). Conversely, an early discovery from basic research may have little value per se, apart from serving as a research tool that may enable further discoveries. The critical incentive challenge is to

---

[5] *Reinganum (1989) provides a comprehensive review of patent races with stochastic innovation. Related work in this setting addresses the role of industry structure on R&D incentives (e.g., Dasgupta and Stiglitz 1980), an issue first addressed formally by Arrow (1962), but which can be traced back to Schumpeter (1942).*

[6] *The ability to license a patented innovation can have other strategic effects on R&D decisions, for example it can be used to discourage imitation and deter entry (Gallini 1984).*

ensure that early innovators are compensated sufficiently for their contribution to later innovators, while also preserving an incentive for the latter to occur. The intertemporal externality of this situation calls for the transfer of profits from successful subsequent applications of a patented innovation to the original inventor(s), which suggests a critical role for licensing (Green and Scotchmer 1995). The problem here highlights the importance of specific features of patents, for example as they relate to patentability and infringement of follow-on innovations. In this setting, the breadth of a patent and its effective life are intimately related. For instance, in a quality ladder context (O'Donoghue, Scotchmer, and Thisse 1998) the economic life of a patent ends with the arrival of the next improvement, the likelihood of which is affected by the novelty requirement and nonobviousness standards set out in the patent statute.

The richer trade-offs that arise in this dynamic context can be illustrated in a setting where two firms (a leader and a follower) compete in a step-by-step innovation contest to develop the dominant product that is marketed at each stage. Here the follower may need to catch up to the knowledge frontier of the leader before having the chance to overtake it. If the follower falls too far behind the leader, it might be rational to drop out of the race, which results in monopoly. One might conjecture that, in this setting, making it easier for the follower to stay in the game by endowing the leader with weak patent rights might be socially desirable. Moschini and Yerokhin (2008) examine the question by focusing on an explicit feature of IPRs, the so-called research exemption. Such a feature is actually extremely limited in the courts' interpretation of U.S. patent law, but it is a central feature of the *sui generis* IPRs established by the U.S. Plant Variety Protection Act (Moschini and Yerokhin 2008). It turns out that the research exemption—which allows followers to use the leader's state-of-the-art knowledge in their research programs—may not be optimal from a welfare perspective. The ranking of the two IPR regimes depends on the relative magnitudes of the costs of initial innovation and improvements, and either regime may dominate from a welfare perspective. The research exemption is most likely to provide inadequate incentives when the cost of establishing a research program is large. An effective research exemption can be seen as limiting the patent right of the leader, consistent with the conjecture that large technological leads are less in need of patent protection than when firms are in neck-and-neck competition. Acemoglu and Akcigit (2012) address this issue of state-dependent patent protection and conclude that, in fact, optimal patents would provide greater protection to more advanced leaders. This trickle-down-of-incentives effect that they uncover ensures that advance leaders still have an incentive to engage in R&D, and also encourages R&D by laggards because of the prospect of greater rewards if and when they become leaders.

Complementarity between innovations similarly affects the incentive effects of patents. This issue is quite relevant in practice because numerous patented innovations are typically required when developing a new product. When patented inputs are highly complementary, or even essential for the new product, each patent holder has blocking power, which makes the production process susceptible to a hold-up problem (Shapiro 2001). The acuteness of this problem is augmented by the fragmentation of the relevant patent rights, which at a minimum increases transaction costs.[7] With respect to the issue of incentives, the existence of complementarities entails an externality effect: a firm pursuing a given innovation exerts a positive externality on another firm that might be pursuing a complementary innovation. This again emphasizes the need for an incentive mechanism that ensures enough profit to induce both innovations and efficient profit-sharing between the complementary innovations, which is very similar to the problem that arises with sequential innovations (which, in fact, is an instance of a specific complementarity relationship).

The incentive justification for patents is challenged by the 1980 Bayh-Dole Act, which allows universities (and businesses) to patent and license (possibly exclusively) discoveries arising from federally sponsored research. The standard incentive rationale for patents can hardly rationalize this Act—if the innovation has already been (publicly) funded and in fact the discovery has been made, allowing it to be patented simply restricts its use (it is equivalent to taxing users). The public pays twice: with the original research grant, and because of supra-competitive pricing of patented products. The alternative of putting the innovation in the public domain should lead to superior welfare outcomes. The Bayh-Dole Act purported to address a different problem, however: its presumption is that without (exclusive) licensing arrangements, firms would not undertake the costly follow-up investment needed to bring an invention to the marketplace. Thus, the motivation is that of promoting "technology transfer" by avoiding the rent dissipation that might prevent commercialization. This act is credited with causing a rapid expansion of university patenting and the widespread creation of university technology transfer offices, although there seems to be little evidence that this rise of university patenting and licensing has significantly affected the level of technology transfer (Sampat 2006).[8] Still, the basic tenet of the Bayh-Dole Act deserves more scrutiny vis-à-vis its stated purpose. The implicit assumption is that the university discovery is patentable, but the follow-up R&D needed for its commercialization is not. Clearly, this presumption does not apply to most products of university research, where downstream R&D may actually have higher patentability as the research moves from basic problems to applied solutions.[9] One can find an incentive-based motive for patenting already paid-for university discoveries, but the argument is more subtle and postulates a "gap between science and the market" characterized by the need to match university inventors with firms that are suitable for carrying out the development stage (Hellman 2007).

---

[7] *In biotechnology this situation has been labeled the "tragedy of the anticommons" (Heller and Eisenberg 1998): the excessive allocation of property rights (i.e., too many gatekeepers with the right to levy a tax) can result in underutilization of the resource (the pool of knowledge). Market-based solutions to this problem include patent pools and cross-licensing (Shapiro 2001).*

[8] *This growth also reflects the universities' seizing on the potential for technology licensing revenue (although boosting such revenues was not, in fact, a motivation for the act). Whereas at present U.S. universities earn about $2 billion per year from such activities, these funds account for a small portion of universities' research expenditures. After accounting for the considerable cost of patenting and licensing, it seems that many U.S. universities are actually net losers, and the positive net returns are highly concentrated in only a few universities (Bulut and Moschini 2009).*

[9] *An important exception, which was actually an element in the political process leading to the act, concerns pharmaceutical drugs, where firms may be reluctant to invest the huge sums required for clinically testing a university-discovered drug unless they can secure an exclusive license.*

Overall Impact of Patents

Even the cursory discussion of the economics of patents provided above suggests that assessing the incentive effects of patents is a complex endeavor. Denicolò (2007) provides a thoughtful analysis that integrates the key insights of the economics of optimal patent design in a stylized model, the parameter of which can be related to the empirical findings of the innovation production function literature. The question is whether patent protection is too high or too low. With the necessary qualifications that apply with such a challenging undertaking, Denicolò (2007) concludes that in the aggregate, patents do not over-compensate innovators. Strong conclusions from more direct empirical analyses on the effectiveness of patents as an incentive mechanism have proven elusive. Hall and Harhoff (2012) review an extensive body of empirical literature and find that there is clear evidence that patents provide effective incentives for innovation only in a few sectors (pharmaceuticals, biotechnology, and medical instruments).

The notion that the incentive effectiveness of patents may vary across sectors of the economy is plausible; indeed, some conclude that patent protection might have adverse effects in some industries. This is possibly the case for the software industry, which is studied in some detail in Bessen and Meurer (2008). These authors' general point is that while the property rights that patents confer do provide private returns to their holders, patents also exert a negative externality on other patent holders who might unwittingly infringe.[10] Taking this risk of litigation into account, their empirical analysis concludes that firms in most sectors (outside the chemical and pharmaceutical industries) would be better off if patents did not exist. An even stronger indictment of IPR institutions is detailed in Boldrin and Levine (2008), who emphasize the need to account for all of the costs of the patent system, in addition to the benefits. These authors document cases of innovative industries that have prospered without strong IPR protection. Moreover, they argue that alternative ways of capturing a return from their investments are anyway available to innovators. They also highlight costs related to the political economy of a patent system, whereby the prospect of patent-based monopoly rights inevitably gives rise to rent seeking, leading to a vicious circle of increased IPR protection that mostly serves the interest of incumbents. The assessment in Boldrin and Levine (2008) is that there is little evidence that IPRs increase innovation; they conclude by advocating the elimination of patents and copyrights.

Dissatisfaction with the actual performance of the patent system is not uncommon. The empirical finding that observed increases in patenting activity in some sectors results from litigation concerns (i.e., a strong patent portfolio can deter competitors from suing, Hall and Ziedonis [2001]), rather than reflecting increased R&D efforts, suggests a strategic role for patenting that is less benign than the purported incentive effect. The expansive interpretation of patentable subject matter, and the apparent failure to hold patent applications to a genuine nonobviousness

---

[10] *Bessen and Meurer's (2008) criticism of the U.S. patent system is organized around the notion of the "notice" function of property (i.e., the ability of the system to notify non-owners of property boundaries), which they argue is problematic for patents because of the intangible nature of the underlying assets.*

standard, might indeed create unnecessary market power positions that lead to inefficiency by limiting access to innovation and deterring further R&D (Jaffe and Lerner 2004), problems that appear serious for the software industry and for business-method and financial patents. All this has led to a number of proposals for reform, but the argument that we would be better off without a patent system remains much more controversial (Ziedonis 2008; Gilbert 2011).[11]

Ultimately, the desirability of patents as incentives for innovation, and the assessment of their efficacy, relates to the specifics of the counterfactual situation. Trade secrets and lead time are oft-mentioned alternatives to patents from the innovating firms' perspective (Coehn, Nelson, and Walsh 2000), but it should be clear that they also entail an *ex post* inefficient provision of the innovation. Indeed, the chief avenue by which innovators can secure a return on R&D is by pricing new products above marginal cost, regardless of what makes this possible. Furthermore, relative to patents, trade secrets have rather different implications vis-à-vis the disclosure of information, diffusion of knowledge, and spillovers of innovation benefits. Lead time (being the first to market) is more benign with respect to the diffusion of information, at least insofar as it is not enabled by trade secrets. But the prospect of lead time might not always be very relevant. Startup firms, for example, require access to capital to proceed to the production stage. In such cases, which are characterized by imperfect and asymmetric information, patents can be a useful signal of innovativeness allowing small innovators and startups to secure financing via venture capital. Patents can also favor the emergence of a market for technology (Gans, Hsu, and Stern 2002).

Contracted Research

Patents, and more generally intellectual property rights, work as incentives to knowledge producers by providing the prospect of a compensation that is tied to the social value of an innovation. Yet, in the case of basic knowledge—a vital input for technological advance—the producers are almost never accorded exclusive legal rights to the fruits of their labors. Instead, the vast majority of basic research is conducted by scientists, engineers, and academicians who operate under a very different set of incentives (Dasgupta and David 2004). Pure ideas are not eligible for patent protection, of course, for a number of reasons. Normatively, assigning property rights to abstract ideas is undesirable because by their very nature they entail a high degree of positive spillovers. Scientists build on each other's ideas, using techniques and concepts invented elsewhere to solve novel problems. Restricting the circulation of ideas and access to pure knowledge would slow down basic research (Aghion et al. 2010). Furthermore, the length of time between the dawn of breakthrough scientific ideas and their commercial application is often very long. In addition to its timing, the eventual commercial application of basic research is marked by a high degree of uncertainty regarding the form it may take, thus making it very difficult for the market to determine *ex ante* the private value of an idea. From a practical point of view, therefore, even if

---

[11] *Such controversies are not new (Machlup and Penrose 1950).*

a scientist could retain the economic rights to an idea, its (private) value would likely be highly discounted and provide a poor innovation incentive.

Given the distinctive attributes of pure knowledge, compensation for basic research has taken alternative forms. As discussed by Dasgupta and David (1994), the economics of incentives literature emphasizes a number of features that specifically apply to the discovery process. First, research is an activity marked by costly monitoring; asymmetric information may render the research sponsor unable to discriminate between research activities in terms of their likelihood of success and expected cost. Research output, however, can be verified by the peer-review process. For this reason, it is common to model research incentives as containing an output-dependent component. Secondly, research is an uncertain process. To the extent that researchers are risk-averse, it is generally suboptimal to have their compensation entirely dependent on the outcome of research. Instead, the research sponsor can generally reduce costs by bearing most of the risk (as the holder of a large portfolio of risky research projects, it has superior diversification opportunities, and in fact it is typically modeled as risk-neutral). Finally, it is often the case that research costs are indivisible, or benefit from economies of scale, which means capital constraints must be taken into account. Typically, resources must be provided to the researcher before the outcome can be observed.

In such cases, the only way to motivate effort is to model the incentives dynamically, so that subsequent support is contingent on the outcome of previous research efforts. In a general form, the problem can be modeled as a dynamic contract where a principal (the research sponsor) tries to maximize the research output of an agent (the researcher) over time, net of the cost of research, by designing a dynamic payment scheme that balances performance incentives with exogenous characteristics, x (e.g., the agent's risk aversion and project characteristics). Typically, models require agents to supply some costly effort, e, that produces a stochastic research output, y, which partly determines payment, w, in subsequent periods. Usually there is also a participation constraint, such that the contract offered by the principal must, in expectation, perform better than some outside option available to the agent.

A tractable version of this approach, where the model is restricted to just one type of research project and two periods, yields a simple and intuitive solution. In such a setting, effort is only supplied in the first period.[12] Given appropriate restrictions on the utility function and the research technology, the total compensation over both periods takes the following simple linear form:

$$w(x, y) = \alpha(x) + \beta(x)y. \tag{1}$$

Holmstrom and Milgrom (1987) present the core result, while Levitt (1995) analyzes the implications for multiple agents when the principal only values the best outcome. Huffman and Just (2000) examine how the values of $\alpha(x)$ and $\beta(x)$ are impacted by different scientist and project characteristics, and they find the outcome-contingent bonus is decreasing

[12] *Because effort is costly and there are no rewards for good outcomes in the terminal period, agents supply a minimum amount of effort in the second period.*

in the researcher's risk aversion, and increasing in a scientist's "talent" for research. They also find that, when a scientist's level of risk aversion is "small" relative to their talent, one's guaranteed salary is decreasing in talent and increasing in risk aversion. Intuitively, if scientists are relatively risk-neutral or are talented enough that the probability of a successful outcome is high, the optimal contract is tightly tied to performance. Conversely, if scientists are highly risk-averse, or if their best efforts still leave much to chance, they must be induced to participate by a high fixed salary. Whereas the linear property in is due to the specific assumptions of the model, the stylized result of a fixed salary plus a bonus that is rising in the current period output is a more general attribute of incentive schemes.

Tournaments and Priority

One manifestation of the compensation mechanism discussed above is provided by tournaments. In a tournament, compensation is tied to the rank order of some measure of output, rather than to the level of output. For instance, scientists receive a guaranteed salary, but supply effort in a bid to have the highest (best or first) output and win a reward in a subsequent period. In this context, one way to model incentives is to assume that per-period research outcomes are transformed into reputational capital. This reputational capital can in turn motivate researchers by appearing directly in the utility function (if scientists care about their social standing, for example), or by appearing as an argument in the $w(x,y)$ function that governs per-period payment. Particularly for scientists operating in the university system, the tournament model is reflected in the priority system, whereby the first scientist (or team of scientists) to publish a discovery is accorded (nearly) all the reputational capital associated with the discovery (Dasgupta and David 1994).

Tournament models of incentives have attractive features. By significantly amplifying the return to small differences in effort and productivity, they produce a strong incentive to supply effort. Furthermore, Levitt (1995) notes that when the link between effort and output is stochastic, there is a sampling benefit to having multiple agents engaged in the same task, as long as the stochastic element is not perfectly correlated across competitors. This is because duplicative research efforts increase the probability that at least one success will be obtained.

There are several specific advantages of the priority system in science (Dasgupta and David 1994). First, priority establishes a reward for successful research that is only activated by disclosure, which allows society to take advantage of the high degree of positive spillover from discoveries. Priority also aligns the social value of a discovery with private value. Society benefits only from the first discovery of an idea—subsequent reinventions and rediscoveries do not add to the stock of usable knowledge—which provides a motivation for rewarding only the first discovery. Furthermore, the priority system also helps address issues of costly verification. Once an idea has been disclosed, it becomes virtually impossible to verify how "close" other scientists were to reaching the same result. By denying rewards to all but the first result, the issue of determining the remaining rank order is evaded. The priority system also induces agents to focus on projects with verifiable qualities, namely those for which there exist people with the appropriate human capital to evaluate a project's contribution.[13]

A problem that might arise in this setting is suboptimal hysteresis. To do research in a field, a scientist must typically accumulate human capital in that field. Thus, a scientist may choose (to continue) to do research in a field because it is populated with scientists who can certify their work, which in turn sustains the supply of scientists in the field, even absent more fundamental justifications. The priority system may also encourage excessive correlation among research strategies (Dasgupta and Maskin 1987). It may, for example, encourage most scientists to pursue a single line of research that appears likely to succeed, but to the neglect of less promising alternatives.[14] While this is individually rational, in the aggregate the likelihood that success will be achieved by at least one scientist is usually improved when researchers pursue diverse strategies. Similarly, the priority system may induce scientists to undertake excessive risk. For example, by focusing on who is first, the priority system may support the allocation of socially excessive resources into reducing the time to discovery by a small amount (days or weeks). This feature may also lead scientists to choose research strategies that have excessively high variance around their expected completion time, since the minimum draw from a high-variance distribution may well be below the minimum of a small variance distribution with a lower mean.

Science careers may increasingly be regarded as a tournament for tenure-track positions. Freeman et al. (2001) argue that a tournament model applies specifically to the life sciences, where a large supply of post-doctoral workers spend several years accumulating reputational capital in the hopes of winning the "prize" of overseeing a lab. Such a description might also apply to other fields where the expansion and globalization of the academic labor market mean the number of PhDs increasingly exceeds the number of highly-desirable tenured positions and high-paying positions in the private sector. [15]

Academic Freedom and Tenure

Researchers often indicate that they were drawn into science because they genuinely take pleasure from solving puzzles, suggesting that research may itself be a consumption good for some scientists. Stern (2004) provides empirical evidence to support this conclusion by exploiting the fact that many postdoctoral biologists receive multiple job offers at the same time, and demonstrates the presence of a trade-off between

---

[13] *Reality is not as stark as theory, and in many cases multiple discoveries made in close succession share credit. One explanation for this may be that it is not, in fact, impossible to distinguish how "close" competitors are, especially when results are published independently and without apparent knowledge of each other. Furthermore, there is value in showing that a result can be replicated.*

[14] *The frequency of "multiples," that is, when a new discovery is made independently by different researchers, also underscores the extent to which correlated research strategies may yield sub-optimal results (Dasgupta and Maskin 1987).*

[15] *Freeman et al. (2001) also worry that the tournament system, at the career level, induces scientists to work too many hours and focus too little on training post-docs and graduate students to be independent researchers. A related concern is that graduate schools do not adequately convey information to prospective students about their prospects in the coming tournament; Freeman et al. (2001), Romer (2001), and Stephan (2012, 162) note that business and law schools may be providing more information about graduate employment than science departments.*

wages and freedom of scientific inquiry. He interprets this as reflecting a taste for research, for which non-basic-research-oriented firms must compensate scientists.[16] The taste-for-research hypothesis is also consistent with evidence that scientists today have a low level of permanent income relative to other human capital intensive professions.[17]

Even if scientists obtain utility from doing their research, the principal-agent problem persists unless their research interests are perfectly aligned with the values of society. One way to exploit the utility that scientists obtain from freedom of inquiry is the tenure system. To the extent that tenured faculty enjoy freedom to pursue their research interests, tenure can be used as a reward for performance, or a prize in the tournament. Furthermore, in addition to providing an incentive for effort, tenure also solves a problem of asymmetric information in research. Carmichael (1988) examines the problem of a university that seeks to assemble the best faculty possible when information about the quality of applicants is restricted to existing department faculty (who have the specific human capital necessary to evaluate the work of peers). Faculty may be able to observe signals about the quality of new hires, but if they recruit a researcher with more talent than themselves they may find themselves replaced. Carmichael shows that the only way for the university to induce faculty to truthfully reveal their signals is to offer them the complete job security carried by tenure.

Competitive Grants

Another mechanism through which scientists enjoy a fixed salary with an output-dependent bonus is the competitive grant system. Grants are typically intended to cover the costs of research, but it is not uncommon to pay a portion of researchers' salaries out of these funds. Indeed, Stephan (2012, 130) reports that, since the 1950s, the responsibility for securing funds for research and buying out of teaching responsibilities has fallen increasingly on university faculty. In extreme cases, faculty may be hired with no guarantee of an income if they fail to bring in a grant. As we have seen above, such high-powered incentives are unlikely to be efficient, given risk-averse faculty, unless research activities have a high probability of success.

Because receipt of the grant is usually essential to conduct the proposed research, grants cannot be used like *ex post* rewards to induce effort. Nor can they be understood in a static context—in a one-shot game setting, there would be no reason for grant recipients to supply effort upon winning a grant. Hence, competitive grants are best modeled dynamically. To gain some insights, consider the simple model outlined in Scotchmer (2004, 249): there exists a population of heterogeneous researchers where, during every period, individuals come up with an idea for a viable research project with probability λ (this also indexes the researcher type), which is unobserved by the granting agency. An applicant who has

---

[16] *For example, he finds that permission to publish research work is associated with about a 20% reduction in wages. Aghion, Dewatripont, and Stein (2008) argue that this may understate the true cost necessary to induce typical scientists to give up their academic freedom, because the sample is restricted to scientists willing to entertain serious job offers from non-science oriented firms.*

[17] *Freeman (2006) finds biologists can expect to earn \$3 million less than medical doctors, and \$1.8 million less than lawyers over their lifetime.*

received a grant must decide whether or not to complete the project at a cost c. Suppose the granting agency follows the simple rule of awarding a grant of value v to every first-time applicant, as well as to every applicant who successfully completed their previous granted project. The awardee's trade-off is between saving c in the current period and being eligible for more grants in the future. Thus, the agent will choose to complete the project when:

$$v \leq v - c + \sum\nolimits_{t=1}^{\infty} \delta^t \lambda(v - c) \tag{2}$$

where δ is the applicant's discount factor. Hence, the condition for the awardee to perform the research work is:

$$\lambda \geq \frac{c(1 - \delta)}{(v - c)}. \tag{3}$$

In equilibrium, therefore, the agency ends up funding only researchers with sufficiently high λ (i.e., those who have ideas most frequently), ensuring that funded projects are actually performed. This illustrates some common features of real-world granting agencies, where the review process typically places considerable weight on past accomplishments (Stephan 2012, 131). Also for this reason, new science hires in academia often receive a "start-up" package, which enables them to accrue successes before applying for grants.

While simple, this approach abstracts away from many elements of the grant game that are important: competition among agents, effort expended to win a grant, the optimal number and sizes of grants to give, and heterogeneity in the value, cost and probability of success of a research project. Lazear (1997) presents a model that incorporates all these elements, except for heterogeneity in potential projects, by representing competition for grants as a raffle. Agents obtain "tickets" as a function of their inherent ability and their efforts (and maybe past accomplishments). For example,

$$T_{it} = \left(e_{ia} + by_{i,t-1}\right)x_i \tag{4}$$

where e denotes application effort, y_{t-1} denotes the quality of last period's outcome, x denotes inherent ability, and b is a parameter indexing the weight placed on previous accomplishments. The probability of winning a grant is then given by the applicant's number of tickets, T_{it}, as a share of the total number of tickets $\Sigma_j T_{jt}$ (thus, an agent's likelihood of winning a grant depends not only on her own effort and ability, but also on the traits of her competitors). An agency with a fixed budget has the problem of deciding how many draws from the "raffle" to make. More draws entail more awards, but at the cost of decreasing the value of a win to applicants. Meanwhile, agents are faced with the problem of deciding how much effort to devote towards winning a grant and, upon receipt, how much effort to spend performing the required work.

Using this framework, Lazear (1997) shows that, in most cases, lower ability and younger applicants will supply more effort. This is because when an applicant is holding few tickets, each additional ticket has a high value (moving from 1 to 2 tickets doubles the chance of winning, for example). High ability people or those with positive values for y_{t-1}, on the other hand, start with a better chance of winning and thus the marginal value of supplying additional effort is low. In a two-period model, older researchers will tend to win more grants.[18] Given that the marginal value of effort provided by a high ability (or more experienced) researcher presumably has higher social value than the marginal effort of low ability types, it may be desirable for the grant system to try to induce highly productive types to exert more effort, which may support the real-world practice of tying the value of a grant to salary.

The tradeoff between number and size of awards can also be examined (Lazear 1997). In general, if fewer grants are awarded, holding the value of each grant constant, researchers will reduce their grant application effort. But if the granting agency has a fixed budget, reducing the number of grants awarded increases the value of each grant. In this setting, it turns out that researchers may exert *more* effort in their applications; although the probability of winning decreases with fewer awards, the increased value of each award may offset this effect and ultimately lead to an increase in effort. Grant-making agencies in the United States experienced this phenomenon in the early 2000s; when the National Institutes of Health (NIH) reduced the number of awards offered after 2002 due to a tightening budget, the number of applications fell. Conversely, when the National Science Foundation (NSF) reduced the number of awards offered between 2000 and 2005, but increased the value of each award, applications per researcher actually increased.[19]

The need for application effort in competitive grant contests suggests that scientists' use of their time may not be optimal: some would argue that scientists' time would be better spent actually doing research. To the extent that a granting agency places weight on y_{t-1} rather than e, this critique is blunted because maximizing the output of an awarded grant effectively becomes an input in next period's grant application. Furthermore, application effort is not necessarily wasteful (e.g., preliminary experiments that have value regardless of the receipt of the grant). Yet some have suggested that grants have come to resemble prizes: applicants submit proposals for work that is already all but complete, ensuring a predictable outcome if the grant is awarded. A related concern is that severe risk aversion now plagues the science grant system (Alberts 2009). Both of these concerns indicate an aspect of grant modeling that until recently had not received much attention: heterogeneity in the riskiness of research.

Research Incentives and Risk-taking

A standard feature of principal-agent models is that the output-contingent portion of researcher compensation should be increasing in output. This seemingly intuitive result, however, is derived from an approach where agents do not choose from a portfolio of possible research projects. Let i label different possible strands of research, and suppose

---

[18] *This conclusion is consistent with stylized facts. For example, the number of researchers under the age of 35 who are awarded the RO1 grant from the NIH has fallen from over 1,000 per annum in 1980 to under 400 per annum by the year 2000 (Freeman and Van Reenen 2009). See also Stephan (2012, 140) for more examples of the disadvantage faced by young researchers.*

[19] *Ironically, the NSF's rationale for offering fewer awards was that it would reduce the time researchers spent on applications (Stephan 2012, 132-133 and 141-143).*

they are characterized by different *ex ante* probabilities of success, p_i, and values of success, y_i. In principle, a risk-neutral granting agency should be indifferent between two projects with the same expected value but different probabilities of success. However, the competitive grant mechanism may engender excess risk aversion. To illustrate this, we return to the model outlined in equations (2) and (3), but suppose that if an agent exerts effort at cost c, there is only a probability $p < 1$ of success. *Ceteris paribus*, this lowers the expected reward of effort, and will encourage shirking. Thus, if the granting agency has knowledge of p, it may reasonably turn down projects with low values of p, since it does not expect a rational agent to complete the grant.[20] Thus, we might expect high-risk grant proposals to be rare. Indeed, there is some evidence that this is the case (Kolata 2009).

A possible strategy to encourage more risk-taking is for the granting agency to shift some of the benefits of diversification onto the scientist by simultaneously financing a set of high-risk projects. By bundling uncorrelated research projects, the probability of uniform failure is reduced for the grant recipient. An obstacle is that such a grant would need to be much larger than typical grants, and would push back the completion time by several periods (assuming it takes longer to run multiple projects). This would tend to increase the incentive to shirk. As a consequence, given the mechanism outlined in equations (2) and (3), only very talented and experienced scientists will tend to be funded (because they stand to gain the most in terms of future grants by completing the projects). Grant-making agencies are aware of these problems, and in response the NIH created the Pioneer Awards in 2004 to specifically fund high-risk but novel and potentially transformative research projects (NIH 2012). At present, these grants represent a small fraction of awards, but they are indeed highly selective, and they provide recipients more funding, time and freedom of research.

To address the relationship between risk-taking and incentives more directly, Manso (2011) explicitly models a researcher who, in a two-period model, has two avenues of research, labeled “exploitative” and “exploratory.” While the value of success is the same regardless of the type of research, there are differences in their probability structure. Exploitative research has a fixed probability of success in both periods, whereas with exploratory research the period two probability depends on the period one outcome. Exploratory research initially has a lower probability but, if the first period is successful, it then has the highest probability of success in period two. The model is meant to capture the trade-off between incremental research that is likely to succeed, but which does not add a great deal to the stock of knowledge, and radical research which may be initially risky, but if successful opens up new arenas for exploration.

When the gains from inducing the exploratory line of research exceed those of the exploitative line, Manso (2011) shows that paying for a good outcome in the first period may actually be detrimental. If the agent is simply compensated for success in each period, and if the principal observes only whether a project was a success or failure (and specifically not whether it was an “exploratory success” or an “exploitative success”),

---

[20] *This may be another argument for tying the value of a grant to salary, assuming that successful risk-takers enjoy a higher salary in subsequent periods.*

it is often optimal for the agent to pursue the "safe" exploitative research in each period. Rewarding success in period one creates an incentive to prioritize the short-term; thus, to induce the agent to explore it is usually best to pay only for success in period two. Indeed, if the probability of success for exploration is very low relative to exploitative research, it may be optimal to pay for failure in period one. This is because a failure in period one signals to the principal that the agent attempted the exploratory research project. To prevent shirking in such a case, compensation in period two must be appropriately designed so that the benefit from succeeding in period two makes exploring a better option than shirking in period one.

While an agent almost never receives a bonus for initial failure,[21] the key message is that compensation schemes should be weighted towards long-term success, and short-term failure should be tolerated if the goal is promoting creative research. There is some evidence that this is the case in the private sector (e.g., Lerner and Wulf 2007). Azoulay, Graff-Zivin, and Manso (2011) tested this theory by comparing the publication record of similar scientists funded either by grants from the NIH or the Howard Hughes Medical Institute (HHMI). Grants from the NIH tend to resemble the grant mechanism discussed earlier, where outcomes in each period are clearly measured and used as inputs in subsequent periods to obtain further grants. HHMI grants, on the other hand, are relatively long-term and very tolerant of early failure. Using propensity-score matching and difference-in-difference approaches, the authors find that HHMI publications tend to be more creative and risky by a number of measures: a higher probability of being a highly-cited article; a tendency to publish with more novel keywords; and a greater diversity of citing journals.

Teamwork

The models considered so far have treated researchers as individuals, a simplification that ignores the fact that scientific work is often conducted by teams. Indeed, the size and frequency of teams has been rising over the last several decades (Wuchty et al. 2007).[22] There are several reasons why this might be the case. Falling communication costs have facilitated collaboration across geographical space, as implied by the finding that papers with authors from multiple universities are rising (Jones, Wuchty, and Uzzi 2008). Indivisibilities and sharing of large datasets and expensive equipment also requires larger teams than in the past. Jones (2009) argues that the rise of teams is an inevitable consequence of humanity's knowledge accumulation; the amount of knowledge needed to contribute to the frontier is now so large that it is spread over several individuals. The rise of teams in research makes the optimality of team incentives an increasingly important issue. Here, we focus on three aspects of incentive design unique to teams: free-riding, sorting, and information sharing.

---

[21] *An interesting exception is from Silicon Valley start-ups, where a capacity to "fail fast" may carry some reputational currency (NPR 2012).*

[22] *Wuchty, Jones, and Uzzi (2007) examine 19.9 million papers and 2.1 million patents between 1955 and 2000 and find that, across major disciplines, both the number of authors of a paper and the frequency of multiple authors have been rising. In science and engineering, for example, the average number of authors rose from 1.9 to 3.5 over the examined period, while the percentage of team-authored papers rose from approximately 50% to 80%.*

The problem of free-riding in teams is addressed by Holmstrom (1982), who shows that an incentive scheme that yields first-best effort from its members can only be achieved if some of the output can be "thrown away" as a punishment to prevent shirking. In this model, members of a team exert a costly effort that contributes to a stochastic output, y, whose distribution is a function of the vector of all individual actions. In this setting, any rule that fully shares the output y between team members will not induce the socially-optimal level of effort—some agents will always have an incentive to shirk because they enjoy the reduction in the cost of effort alone, but the reduction in the expected reward is split between all the team members. However, it may be possible to get arbitrarily close to first-best effort levels if the following sharing rule is used (Holmstrom 1982):

$$s_i(y) = \begin{cases} s_i y, & y \geq \bar{y} \\ s_i y - k_i, & y < \bar{y} \end{cases} \tag{5}$$

where s_i denotes agent i's share (so that $\sum_i s_i = 1$). Note that if output falls below some $\bar{y}$, then total team compensation falls below its output level y. For appropriately-chosen values of $\bar{y}$ and k_i, this rule can induce optimal effort. Some incentive schemes in research may approximate this result. In academic publishing, for example, if a paper's quality, y, falls below a threshold it is not accepted for publication, so that the reputational capital awarded to each and all of the authors for this work is essentially nil (i.e., $k_i \approx s_i y$).[23]

The necessity of teams also introduces the issue of optimal matching, i.e., the optimal grouping of individuals of different abilities. This question is addressed by Kremer (1993), whose O-ring production model in its simplest form is:

$$E(y) = \prod_{i=1}^{n} q_i \tag{6}$$

where n tasks must be completed by n agents who vary by ability, $q_i \in [0,1]$. In such models, assortative matching takes place so that individuals with similar abilities work together. For example, academic departments tend to hire faculty of similar ability levels; to the extent that they are the units of collaboration, this is consistent with assortative matching. As the cost of communication falls, collaboration across universities has become increasingly common. Nevertheless, in an analysis of multi-university collaborations, Jones, Wuchty, and Uzzi (2008) find that collaborations between schools that are ranked similarly are more common than predicted by a random matching model, and that this trend has risen over time.

When a team structure is necessary for innovation, issues of information-sharing are relevant. Niehaus (2011) develops a framework to study the costs of sharing information in a setting where agents share knowledge with others whenever the private benefits exceed communication costs. If completing a project requires a common set of information for

---

[23] *If there are constraints on the maximum value of k_i(e.g., $s_i(y)$ cannot be negative), then there may not exist a value of $\bar{y}$ that can induce first-best effort. Holmstrom (1982) argues that this implies a maximum team size that can be effectively managed by such an incentive scheme.*

all team members, then the cost of sharing information must be considered a part of the project's cost. In a multi-disciplinary setting, this cost can be substantial. This may make such projects unviable when considered individually, but worthwhile if a set of projects that require the same common set of information can be found.[24] Whether such fixed set-up cost arguments are sufficient to explain the desire of many schools to establish interdisciplinary programs is an open question. Alternatively, the currently widespread enthusiasm for interdisciplinary projects may grossly underestimate their true costs.

A distinct information-sharing situation arises when two agents pursue independent projects, but have the opportunity to learn from each other. This is the context analyzed by Ederer (2010), who specifically considers a variation of Manso (2011) where each researcher has an exploitative and exploratory research path open to her. Again, success with the exploratory method in period one means a higher expected probability of success in period two. However, here an agent can observe whether his co-researcher succeeded or failed with the exploratory method, and in doing so gain the benefit of exploratory success for period two. It follows that agents have an incentive to use the exploitative method and let their co-researcher bear the risk of the exploratory method. Ederer (2010) shows that in such a setting, an optimal compensation scheme depends on the joint outcome, rather than the individual's, so that researchers are rewarded for the success of their co-researcher as well as their own output.

Prizes

The two broad mechanisms for securing innovation discussed above are problematic, though for different reasons: patents impose a monopoly deadweight loss, while contracted research must overcome moral hazard problems. An alternative that can potentially overcome these limitations is the innovation prize: an *ex post* reward for innovations meeting certain prerequisites. If the prize is appropriate, it may generate enough investment/effort to bring about a desired innovation. Having compensated the inventor, the innovation can then be put in the public domain, thereby eschewing the deadweight welfare loss associated with patents.

Following Shavell and van Ypersele (2001), consider an innovation that can be realized with probability $P(c)$, at unobservable cost c. This probability function is increasing in c and satisfies $P'(c) > 0$ and $P''(c) < 0$. An individual determining how much to invest in the discovery solves:

$$\max_{c} \ BP(c) - c \tag{6}$$

where B is the benefit (value) to the individual of a successful discovery. Note that if the true social value is V, then the social planner would want to have $B = V$, so as to induce the individual decision-maker to choose the socially optimal level of investment c^*, which satisfies $VP'(c^*) = 1$. This outcome can be achieved if a prize of value V is offered to the successful discovery, since this forces the researcher to solve the social planner's problem (it is assumed that, upon discovery, the innovation is put in the

---

[24] *It is also costly to verify the quality of interdisciplinary work if the evaluators must possess information from a range of disciplines, though this cost is not directly borne by the authors.*

public domain and is supplied at marginal cost by a competitive market). If social welfare is indifferent to income distribution so that lump-transfers do not reduce welfare, prizes can be a first-best solution to the innovation problem. Note that a patent system cannot achieve this outcome. Under a patent system the individual reward is the monopoly value of the innovation π, so that the optimal solution $\hat{c}$ satisfies $\pi P'(\hat{c}) = 1$. Because $P(c)$ is concave, it follows that $\hat{c} < c^*$ (inefficiently low investment). Furthermore, as noted earlier, the patent will impose an *ex post* deadweight loss by restricting the use of innovation (including possible hold up or blocking of subsequent innovations).

When Prizes Are Second-best

At least three considerations reduce the attractiveness of prizes. First, as noted by Wright (1983) in his seminal examination of the trade-off between alternative forms of innovation incentives, it is generally not the case that only one firm responds to the prize incentive. Assume instead that there are N firms capable of undertaking the investment and that, if multiple firms succeed, the prize is split between them. In such a setting, it is highly unlikely that aggregate investment and the implied probability of at least one success are socially optimal. In fact, Wright (1983) argues that this "common pool" effect is likely to induce excessive aggregate investment.[25] A second drawback of prizes, discussed by Gallini and Scotchmer (2002), relates to the assumption of lump-sum transfers. If frictionless transfers are not possible, then prizes will also induce some distortions in raising the necessary funds.[26] Finally, the assumption that a new discovery will instantly become available to all, driving profits to zero when prizes replace patents, may neglect the fact that competitive advantages associated with the innovation process (opportunity of first-mover advantages, secrecy on some aspects of the discovery not needed to claim the prize) may still provide profit opportunities for the innovator. In such a case, deadweight losses may be unavoidable, unless prizes are coupled with a legal requirement to sell at marginal cost (which would have its own verification costs).

When Innovators Know More Than the Planner

The drawbacks discussed above diminish the achievable efficiency gains that can be had from a prize system, but they generally apply to patents as well. The patent system, however, looks more attractive once considerations of asymmetric information are taken into account (Wright 1983). This is related to the difficulty of evaluating the prospects for, and output of, research. It seems plausible that innovators have better information about the costs of research, the probability of success, or the value of an innovation.

When the cost of research is unknown to the social planner but known to the innovator, in the model discussed earlier (and ignoring the issue of

---

[25] *Whereas the implication here is that the socially optimal level of investment may be unattainable with prizes, other winner-takes-all reward systems (such as patents) suffer the same defect.*

[26] *The belief is usually proffered that such distortions are likely less significant than the deadweight loss imposed by patents (the argument is that monopoly pricing due to a patent effectively imposes a tax in a single market, while revenue for funding prizes can in principle be raised with an optimal taxation scheme that spreads distortions over many markets).*

multiple firms competing for a prize) the optimal prize is unaffected and should remain equal to V. Such a prize continues to induce the socially optimal level of investment. If, alternatively, one models innovators as enduring a fixed cost, K, to achieve an innovation of value, V_i (e.g., Weyl and Tirol 2012; Chari, Golosov, and Tsynniski 2010), then it is sufficient to offer prizes that compensate innovators for K. In fact, if raising revenue to fund prizes induces distortions, then it is desirable to keep prizes at the minimum level to induce socially valuable innovations (i.e., those with $V_i \geq K$). In such a scenario, if costs were unknown, the benefits of a prize system would be reduced. Note, however, that in such a case prizes continue to outperform patents. Because we assumed that the value of innovations is known, it remains possible to offer prizes equal to the value of a patent. This will induce the same innovations that would be induced under a patent system, but avoid the associated monopoly deadweight loss. And, as shown by Wright (1983), unknown costs make prizes (as well as patents) superior to contracting research.

The other case of interest is when the innovator is better able than the planner to estimate the social value of the innovation (value V depends on consumer demand for the innovated product, and private firms may well have superior information about such markets). This presents a more pernicious problem than when only the cost of innovation is in doubt. In the Shavell and van Ypersele (2001) model outlined earlier, the value, V, is a random variable (a function of the distribution of demand functions). It is assumed that all innovations are valuable, and prizes are frictionless transfers. Under a patent system, the reward to innovation, π, rises in tandem with the true social value, V, so that more resources are expended on the production of the most valued innovations. In contrast, under a prize system, because it is impossible to distinguish the value of one innovation from another, all are awarded the same prize. The problem of the planner is to pick this amount, and it turns out that the optimal prize is simply the expected innovation value, $E(V)$.

Provided that lump-sum transfers are possible, an advantage of prizes is that they do not induce deadweight loss. Instead, the loss comes only from distortions in investment. The innovators' optimality condition is $E(V)P^{'}(c^*) = 1$, and thus innovators supply the same investment to all innovations: excess investment for innovations that have lower than average social value, and insufficient investment for innovations that have higher than average social value. Thus, prizes will tend to perform better than patents when the distribution of V has a narrow variance (implying that $E(V)$ is a good proxy for the true V), and worse than patents when the variance is large.

In any case, a hybrid system where innovators can choose to receive either a patent or a prize dominates a pure patent system. To see why, suppose the distribution of V is bounded from above and below by $\bar{V}$ and $\underline{V}$, respectively, and that π is similarly bounded by $\bar{\pi}$ and $\underline{\pi}$. Note that a pure patent system is indistinguishable from a hybrid system where the prize is worth less than $\underline{\pi}$ (in such a situation the value of a patent always exceeds that of the prize). Now consider raising the prize to $\underline{V}$. For the lowest value innovation, it is optimal for the innovator to choose the prize, and the prize induces the socially optimal investment while eliminating the associated deadweight loss. The prize is also chosen for all innovations such that $\underline{V}$ is greater than the patent value. Since $\underline{V}$ is the lower bound on values, the prize will never induce excess investment, but will instead bring investment levels closer to their optimum (relative to patents) and remove deadweight loss. Such a hybrid system thus dominates a pure patent system.[27]

Under a hybrid system, inefficiencies arise only from the deadweight loss on high-value patented innovations, and from incorrect levels of investment (too little for high value and too much for low value). A second potential drawback of a hybrid prize system would be overpaying for low value innovations—one can imagine the emergence of a class of rent-seeking innovators who "discover" products of zero value and receive prizes. Chari, Golosov, and Tsynniski (2010) consider a model that allows firms to invest zero in return for an innovation with zero social value. Because the planner has no way to verify the value of the innovation, nor the cost endured in its discovery, any prize of a positive value presents a riskless profit opportunity. The conclusion is that if raising revenue to fund prizes introduces economic distortions, if the planner has no way to verify the value of innovations, and if it is possible to cheaply create worthless innovations, then the deadweight loss associated with raising revenue for prizes to low-value innovations may make a prize system (or even a hybrid system) just as inefficient as a patent system.

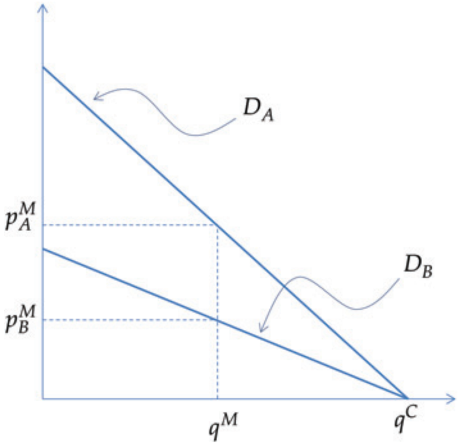
Information, the Market, and Collusion

The assumption that the planner knows only the distribution of innovation values is restrictive—if prizes can be given after the innovation is brought to market, sales of the product can provide (*ex post*) information about the true value of an innovation. In Shavell and van Ypersele (2001), prizes determined after observing the quantity sold at a marginal cost will tend to be closer to the true value, and will therefore better approximate the social optimum. Weyl and Tirole (2012), however, point out that assessing the value of an innovation from the quantity sold ignores an additional relevant source of heterogeneity, namely the consumers' willingness to pay. The crux of the matter is illustrated in figure 2, which reports the demand function of two different goods. When sold at marginal cost (here normalized to zero), goods A and B have the same demand q^C; yet, it is apparent that good A is valued more (consumers enjoy a larger surplus with good A). The reward provided by a patent is responsive to a consumer's willingness to pay because the price charged by the innovator-monopolist reflects the nature of demand. In figure 2, the monopoly prices p_A^M and p_B^M would reveal the fact that good A carries a higher value. The sorting role of market power tilts the comparison of prizes and patents in favor of the latter. More generally, the analysis shows that there is a tradeoff between the value of extra information obtained from non-marginal cost pricing, since it could be used to design more efficient prizes, and the cost of obtaining this information, in the form of deadweight loss. Weyl and Tirole (2012) explore this tradeoff by means of a reward function $t(p,q)$ that depends on both the price and the quantity charged by the innovator. For example, the Cobb-Douglas reward function $t(p^a q^{1-a})$ would yield a patent-like reward for $a = 1/2$, and a

---

[27] *More generally, Shavell and van Ypersele (2001) show that, under a hybrid system, one can do even better by choosing a prize that exceeds the lower bound.*

Figure 2 Market Power Identifies Higher-Value Innovations



quantity-based prize for $a = 0$. They find that, in general, the optimal reward policy entails $a \in [0, 1/2]$, i.e., a patent/prize hybrid.

The objective of conditioning innovators' rewards on market values while still reducing deadweight losses also motivates Kremer (1998), who exploits the fact that competitors are likely to observe signals about an innovation's value. In his patent-buyout mechanism, an innovator receives a patent, but the rights to this patent are auctioned off. The presumption is that competitors' bids will reveal information about the patent's private value. The planner then uses this information to estimate the social value of the innovation,[28] and offers a corresponding reward to the innovator in exchange for giving up the patent rights. If the innovator consents, the planner pays the patent holder and then makes the innovation freely available with probability $(1 - p)$, and sells it to the highest bidder for the second highest bid with probability p. While p may be quite small, so that most innovations are in fact made freely available, it is necessary to give auction participants an incentive to bid honestly by providing a non-negligible chance they will really have to pay for the patent.

The obvious potential virtues of the foregoing patent-buyout mechanism critically depend on the assumption that competitors do have good information, and that there are enough such firms for the auction to work effectively. More generally, whenever the planner uses signals to determine the size of a prize, there may be incentives for firms to manipulate these signals. An innovator could instruct his competitors to bid above their valuation, promising to pay them the difference between their true valuation π and their bid if they are forced to actually purchase the patent. Kremer (1998) was acutely aware of this problem and suggested a number of strategies to raise the costs of collusion. Chari, Golosov, and Tsynniski (2010) explicitly derive an optimal policy when collusion is possible, and obtain a series of stark results depending on information available to the planner and the potential for collusion. In their model, it is possible to create innovations with zero value for zero cost. When the planner cannot observe the value of innovations, it is optimal to use a patent system, because otherwise firms can make riskless profits by

---

[28] *Kremer (1998) suggests marking up the third-highest bid by a predetermined proportion.*

submitting zero-value innovations for prizes. However, when competitors observe a signal and the government can consult with them, or when the government can observe the number of units sold, it becomes optimal to use a pure prize system. Conversely, when these signals can be manipulated, so that firms may bribe each other to report inflated signals, patents once again become the optimal instrument.[29]

Prizes in Practice

The preceding review suggests that prizes may in principle be superior to patents, but only if it is easy to estimate the true value of an innovation with a relatively high degree of certainty, and if market information is costly to manipulate. However, relevant tradeoffs between a patent system and a prize system are difficult to gauge because a large prize system on a scale with a national patent system has never been tried. Several questions arise. For example, how costly would it be to estimate the appropriate size of innovations, especially compared to the cost of administering the patent system? Would innovators challenge prize decisions in costly legal battles? How difficult would it be to manipulate signals about an innovation's quality? Would we see a surge in research spending as innovators capture a greater share of consumer surplus than they do under the patent system? These questions are difficult to answer. A number of historical examples are tantalizing, but arguably insufficient for drawing strong conclusions. Kremer (1998) discusses the French government's purchase of the patent for Daguerreotype photography in 1839, and the U.S. government's purchase of the cotton gin patent in 1802, both of which appeared to spur rapid adoption and technical improvements of the respective technologies. However, the famous case of the prize for a solution to the longitude problem advertised in 1714 by the British government, discussed by Kremer and Williams (2010), illustrates a counter-example. The inventor John Harrison was never paid the full amount of the prize because the prize committee was unconvinced of his solution's accuracy—an illustration of the inherent difficulty for a planner to assess the true value of an innovation.

Other case studies have focused on the performance of prizes contests, operated in parallel with patent systems, in the more distant past.[30] Brunt, Lerner, and Nicholas (2011) examine annual prize competitions conducted by the Royal Agricultural Society of England between 1839 and 1939, while Nicholas (forthcoming) looks to the role of 8,503 prize contests held in Japan between 1886 and 1911. Moser and Nicholas (2012) examine the role of prizes at the 1851 World Fair in U.S. innovation. Typically these studies link prizes won or offered in various technological subclasses to patent output in the same subclass, and thereby demonstrate how prizes

---

[29] *Chari, Golosov, and Tsynniski (2010) show that if it is costly to falsify signals, an intermediate result is obtained. For example, firms could try to inflate the apparent demand for their products by purchasing goods themselves. If costs are high, a pure prize system may be optimal; otherwise patents must be part of an optimal solution.*

[30] *Yet another strand of work has investigated more particular questions. Boudreau and Lakhani (2012) perform an experiment where computer programmers compete to solve problems in exchange for a prize, but primarily examine the impact of self-sorting on worker effort. Boudreau, Lacetera, and Lakhani (2011) analyze the results of a large sample of computer programming contests to see the impact of the size of the competition pool on effort and the likelihood of an extreme-value solution.*

indeed induce innovation. While these instances show the power of prize competitions to induce innovation, such prizes operated in a manner quite different from the theoretical models discussed earlier. Most notably, in the majority of cases, prizes took the form of medals, rather than monetary rewards that attempted to compensate an inventor for costs or the social value of an innovation. These prizes likely induced innovation through channels that were distinct from the pure incentives effect typically of interest. In particular, such prizes signaled quality and communicated areas of technological opportunity, two important functions in an era when communication costs were high.

In the absence of solid evidence about how a modern large-scale prize scheme would operate, Kremer and Williams (2010) advocate increased experimentation with so-called voluntary prizes applied to sectors of particular promise (voluntary so that innovators do not hold back investment for fear the patent system will be abandoned). Medical advances seem particularly amenable to prizes, since their social value can be approximated reasonably well (by multiplying the dollar value of a healthy year of life by the expected years of life saved). Prizes in this sector include Advanced Market Commitment (AMC) prizes, which guarantee a per-unit price to be paid by the prize's sponsors for vaccines that target diseases in poor countries. Williams (2012) reports that an analysis for an AMC prize is currently underway, whereby the efficacy of the prize in inducing discoveries for pneumococcal vaccines will be compared with a reference group composed of vaccines for diseases with similar characteristics but which are ineligible for the prize. Another proposal is the Medical Innovation Prize fund, which proposes devoting 0.5% of U.S. GDP towards a fund that will pay for innovation across a wide range of medical advances.

Conclusions

Unprecedented economic growth rates since the dawn of the industrial revolution have led to drastically improved standards of living for a vast and increasing portion of the world's population. This growth has also brought new problems (e.g., rising inequality, depletion of nonrenewable resources, environmental impacts possibly as serious as global climate change) that highlight the challenges of supporting continued economic development. Some argue that the very presumption of continued growth is unrealistic (Gordon 2012). Opinions differ on this controversial issue, but there is consensus that, if growth is to be sustained, one element is essential—continued innovation.

What is required to foster innovation is a difficult and multifaceted question. Without intending to trivialize this complex issue, in this paper we have taken the view that the role of incentives is critical. Some recent work, neatly summarized in Acemoglu and Robinson (2012), emphasizes that political and economic institutions are decisive for development and growth. Benign ("inclusive") institutions—which, at the macro level considered by the authors, encompass property rights, contract enforcement, freedom of individuals to choose their economic activity, and generally competitive markets—work by providing the incentives needed to unlock individual creativity and investment, thereby engineering a sequence of

Schumpeterian "creative destruction" such that good products and processes are replaced by better ones in a wave of continuous innovation.

At the more specific and concrete level of existing R&D institutions and policies, recognition of the crucial role played by incentives still leaves a number of open questions. Scientific research and technological development span a range of heterogeneous activities where a number of (imperfect) incentives systems are at work. In this paper we have reviewed key innovation incentive mechanisms, with particular attention to more recent contributions and an eye towards distilling what consensus there might be to inform the economic discussion of current and emerging policy issues.

Intellectual property rights remain one of the most important innovation-oriented institutions in modern economies. This is particularly the case in an environment where the largest share of R&D activities is carried out by industry, and where it is unlikely that significant additional public funds might be mobilized to support R&D. Our discussion of patents has emphasized their potentially unique role in fostering innovation in a decentralized market economy. However, the idealized function of patents appears to be rather imperfectly implemented in real-world settings. Cumulative innovations, for example, raise unresolved issues regarding how IPRs can best facilitate innovation. In particular, the potential negative effects of IPRs becoming an obstacle for future innovations has been recognized, as have the political economy implications of perverse incentives arising with the pursuit of any kind of monopoly rights. Bad patents (e.g., for innovations that do not meet genuine non-obviousness standards) are all too common. Easy and rather obvious improvements to existing technologies, whether new or not, hardly deserve exclusive rights for a long period. Some fast-moving areas of the economy, such as the software industry underpinning the information and communication sector, appear particularly ill-served by the current patent system. There are many ripe avenues for reform, and the vast body of economic analyses, some of which we have highlighted in this paper, can potentially be very valuable for improving the patent system.

The alternative of innovation prizes is theoretically alluring—witness the resurgence of interest that we briefly reviewed in this paper—but ultimately seems to be best-suited to a narrow set of applications. In the (perhaps rare) case when the government has the ability to estimate an innovation's value with a high degree of confidence, prizes would seem to be ideal, as they combine the *ex ante* inducement power of a patent system (the prospect of a sizeable reward) with the absence of *ex post* deadweight losses. For most potential innovations, however, firms possess special information about what innovations are actually possible, as well as their likely market value, and might otherwise have the ability to manipulate a government's estimates of value under a prize system. In such cases a patent system, which sacrifices some *ex post* efficiency in the form of deadweight loss and potential hold-up costs, can nevertheless align private investment in innovations with true social value. Furthermore, even when they are legitimate, innovation prizes may not be scalable, as they would require considerable additional R&D funds from public sources, which is an unattractive proposition in the current fiscal climate characterized by concerns over the government's budget deficit and the size of public debt.

Direct governmental financing of research through a system of dynamic incentives may, in some cases, be the best system for generating new knowledge. This is so when the value of a discovery is highly uncertain, even to its discoverers, and when the degree of positive spillovers becomes very high (such that a patent system may prove insufficient to spur the desired R&D investment). Much of the presently conducted basic research is directly supported by public funds, and it is difficult to see how it could be done otherwise. Indeed, in an era of scarce public resources, the argument could be made that public funds should concentrate even more on basic science devoted to discoveries that potentially have a broad set of applications and large long-run payoffs. Such direct procurement of knowledge by the government must rely on disbursement mechanisms that provide suitable incentives, given the agency issues that inevitably arise. In this setting, an emerging concern is that grant-making agencies have become so risk-averse that potentially transformative research agendas are left to stagnate. Changes to the standard system for allocating competitive grants to encourage increased risk-taking, and an expanded focus on more radical research, might be warranted. Increasing support for individual researchers, as opposed to their projects, is a possible way forward. The challenge here is to find workable ways to do so that eschew the dangers of researchers' opportunistic behavior under moral hazard and asymmetric information. Some of the economic work that we have reviewed and discussed in this paper provides useful and promising insights in these directions.

Acknowledgements

Matthew S. Clancy was supported by a USDA National Needs Fellowship grant. GianCarlo Moschini was supported by a USDA Policy Research Center grant to the Center for Agricultural and Rural Development and by the endowment of the Pioneer Chair in Science and Technology Policy. The authors thank the journal's anonymous reviewers for constructive comments.

References

Acemoglu, Daron, and Ufuk Akcigit. 2012. Intellectual Property Rights Policy, Competition and Innovation. *Journal of the European Economic Association* 10 (1): 1–42.

Acemoglu, Daron, and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. New York: Crown Publishers.

Aghion, Philippe, Paul A. David, and Dominique Foray. 2009. Science, Technology and Innovation for Economic Growth: Linking Policy Research and Practice in 'STIG' Systems. *Research Policy* 38 (4): 681–693.

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein. 2008. Academic Freedom, Private-Sector Focus, and the Process of Innovation. *RAND Journal of Economics* 39 (3): 617–635.

Aghion, Philippe, Mathias Dewatripont, Julian Kolev, Fiona Murray, and Scott Stern. 2010. The Public and Private Sectors in the Process of Innovation: Theory and Evidence from the Mouse Genetics Revolution. *American Economic Review: Papers & Proceedings* 100: 153–158.

Alberts, Bruce. 2009. On Incentives for Innovation. *Science* 326 (5957): 1163.

Arrow, Kenneth J. 1962. Economic Welfare and the Allocation of Resources for Inventions. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. Richard R. Nelson, 609–626. Princeton, NJ: Princeton University Press.

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. 2011. Incentives and Creativity: Evidence from the Academic Life Sciences. *RAND Journal of Economics* 42 (3): 527–554.

Bessen, James, and Eric Maskin. 2009. Sequential Innovation, Patents, and Imitation. *RAND Journal of Economics* 40 (4): 611–635.

Bessen, James, and Michael J. Meurer. 2008. *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk.* Princeton, NJ: Princeton University Press.

Boldrin, Michele, and David K. Levine. 2008. *Against Intellectual Monopoly.* Cambridge: Cambridge University Press.

Boudreau, Kevin J., Nicola Lacetera, and Karim R. Lakhani. 2011. Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. *Management Science* 57: 843–863.

Boudreau, Kevin J., and Karim R. Lakhani. 2012. The Confederacy of Heterogeneous Software Organizations and Heterogeneous Developers: Field Experimental Evidence on Sorting and Worker Effort. In *The Rate and Direction of Inventive Activity Revisited,* ed. Josh Lerner and Scott Stern, 483–505. Chicago: University of Chicago Press.

Brunt, Liam, Josh Lerner, and Tom Nicholas. 2011. Inducement Prizes and Innovation. *Harvard Business School Working Paper* 11–118.

Bulut, Harun, and GianCarlo Moschini. 2009. U.S. Universities' Net Returns from Patenting and Licensing: A Quantile Regression Analysis. *Economics of Innovation and New Technology* 18 (2): 123–137.

Carmichael, Lorne. 1988. Incentives in Academia: Why is There Tenure? *Journal of Political Economy* 96 (3): 453–472.

Chari, Varadarajan V., Mikhail Golosov, and Aleh Tsynniski. 2010. Patents and Prizes: Using Market Signals to Provide Incentives for Innovation. *Journal of Economic Theory* 147 (2): 781–801.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. 2000. *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)*. National Bureau of Economic Research Working Paper 7552.

Dasgupta, Partha, and Eric Maskin. 1987. The Simple Economics of Research Portfolios. *Economic Journal* 97 (387): 581–595.

Dasgupta, Partha, and Paul A. David. 1994. Toward A New Economics of Science. *Research Policy* 23: 487–521.

Dasgupta, Partha, and Joseph Stiglitz. 1980. Uncertainty, Industrial Structure and the Speed of R&D. *Bell Journal of Economics* 11 (1): 1–28.

Denicolò, Vincenzo. 2007. Do Patents Over-Compensate Innovators? *Economic Policy* 22: 679–729.

——. 1996. Patent Race and the Optimal Patent Breadth and Length. *Journal of Industrial Economics* XLIV (2): 249–265.

Ederer, Florian. 2010. Launching a Thousand Ships: Incentives for Parallel Innovation. Working Paper, School of Management, University of California, Los Angeles.

Freeman, Richard. 2006. Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership? In *Innovation Policy and the Economy, Volume 6*, ed. Adam B. Jaffe, Josh Lerner, and Scott Stern, 123–157. Cambridge, MA: MIT Press.

Freeman, Richard, Eric Weinstein, Elizabeth Marincola, Janet Rosenbaum, and Frank Solomon. 2001. Competition and Careers in Biosciences. *Science* 294 (5550): 2293–2294.

Freeman, Richard, and John van Reenen. 2009. What if Congress Doubled R&D Spending on the Physical Sciences? In *Innovation Policy and the Economy, Volume 9*, ed. Josh Lerner, and Scott Stern, 1–38. Chicago: University of Chicago Press.

Gallini, Nancy. 1984. Deterrence by Market Sharing: A Strategic Incentive for Licensing. *American Economic Review* 74 (5): 931–41.

———. 1992. Patent Policy and Costly Imitation. *RAND Journal of Economics* 23: 52–63.

Gallini, Nancy, and Suzanne Scotchmer. 2002. Intellectual Property: When Is It the Best Incentive System? In *Innovation Policy and the Economy, Volume 2*, ed. Adam B. Jaffe, Josh Lerner, and Scott Stern, 51–77. Cambridge, MA: MIT Press.

Gans, Joshua S., David H. Hsu, and Scott Stern. 2002. When Does Start-Up Innovation Spur the Gale of Creative Destruction? *RAND Journal of Economics* 33 (4): 571–586.

Gilbert, Richard. 2011. A World Without Intellectual Property? A Review of Michele Boldrin and David Levine's Against Intellectual Monopoly. *Journal of Economic Literature* 49 (2): 421–432.

Gilbert, Richard, and Carl Shapiro. 1990. Optimal Patent Length and Breadth. *RAND Journal of Economics* 21 (1): 106–112.

Gordon, Robert J. 2012. Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds. National Bureau of Economic Research Working Paper 18315.

Grabowski, Henry G., Margaret Kyle, Richard Mortimer, Genia Long, and Noam Kirson. 2011. Evolving Brand-Name and Generic Drug Competition May Warrant a Revision of the Hatch-Waxman Act. *Health Affairs* 30 (11): 2157–2166.

Green, Jerry R., and Suzanne Scotchmer. 1995. On the Division of Profit in Sequential Innovation. *RAND Journal of Economics* 26 (1): 20–33.

Hall, Bronwyn H., and Dietmar Harhoff. 2012. Recent Research on the Economics of Patents. *Annual Review of Economics* 4: 541–565.

Hall, Brownwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. National Bureau of Economic Research Working Paper 8498.

Hall, Bronwyn H., and Rosemarie Ham Ziedonis. 2001. The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995. *RAND Journal of Economics* 32 (1): 101–28.

Heller, Michael A., and Rebecca S. Eisenberg. 1998. Can Patents Deter Innovations? The Anticommons in Biomedical Research. *Science* 280 (5364): 698–701.

Hellmann, Thomas. 2007. The Role of Patents for Bridging the Science to Market Gap. *Journal of Economic Behavior & Organization* 63 (4): 624–647.

Holmstrom, Bengt. 1982. Moral Hazard in Teams. *Bell Journal of Economics* 13 (2): 324–340.

Holmstrom, Bengt, and Paul Milgrom. 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* 55 (2): 303–328.

Huffman, Wallace E., and Richard E. Just. 2000. Setting Efficient Incentives for Agricultural Research: Lessons from Principal-Agent Theory. *American Journal of Agricultural Economics* 82 (4): 828–841.

Jaffe, Adam B., and Josh Lerner. 2004. *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It*. Princeton, NJ: Princeton University Press.

Jones, Benjamin F. 2009. The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *Review of Economic Studies* 76 (1): 283–317.

Jones, Benjamin F., Stefan Wuchty, and Brian Uzzi. 2008. Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. *Science* 322 (5905): 1259–1262.

Klemperer, Paul. 1990. How Broad Should the Scope of Patent Protection Be? *RAND Journal of Economics* 21 (1): 113–130.

Kolata, Gina. 2009. Forty Years War: Grant System Leads Researchers to Play It Safe. *The New York Times*, June 27. http://www.nytimes.com/2009/06/28/health/research/28cancer.html?_r=2 (accessed on 15 October 2012).

Kremer, Michael. 1993. The O-Ring Theory of Economic Development. *Quarterly Journal of Economics* 108 (3): 551–575.

——. 1998. Patent Buyouts: A Mechanism for Encouraging Innovation. *Quarterly Journal of Economics* 113 (4): 1137–1167.

Kremer, Michael, and Heidi Williams. 2010. Incentivizing Innovation: Adding to the Tool Kit. In *Innovation Policy and the Economy, Volume 10*, ed. Josh Lerner and Scott Stern, 1–17. Chicago: University of Chicago Press.

Langinier, Corinne, and GianCarlo Moschini. 2002. The Economics of Patents. In *Intellectual Property Rights in Animal Breeding and Genetics*, ed. Max F. Rothschild and Scott Newman, 31–50. New York: CABI Publishing.

Lazear, Edward P. 1997. Incentives in Basic Research. *Journal of Labor Economics* 15 (1): S167–S197.

Levitt, Steven D. 1995. Optimal Incentive Schemes When Only the Agents "Best" Output Matters to the Principal. *RAND Journal of Economics* 26 (4): 744–760.

Machlup, Fritz, and Edith Penrose. 1950. The Patent Controversy in the Nineteenth Century. *Journal of Economic History* 10 (May): 1–29.

Manso, Gustavo. 2011. Motivating Innovation. *Journal of Finance* 66 (5): 1823–1860.

Moschini, GianCarlo. 2004. Intellectual Property Rights and the World Trade Organization: Retrospect and Prospects. In *Agricultural Policy Reform and the WTO,* ed. Giovanni Anania, Mary E. Bohman, Colin A. Carter, and Alex F. McCalla, 474–511. Aldershot, UK: Edward Elgar Publishing.

Moschini, GianCarlo, and Oleg Yerokhin. 2008. Patents, Research Exemption, and the Incentive for Sequential Innovation. *Journal of Economics & Management Strategy* 17 (2): 379–412.

Moser, Petra, and Tom Nicholas. 2012. Prizes and the Publication of Ideas. Working Paper 13-031, Harvard Business School.

Nicholas, Tom. Forthcoming. Hybrid Innovation in Meiji, Japan. *International Economic Review*.

Niehaus, Paul. 2011. Filtered Social Learning. *Journal of Political Economy* 119 (4): 686–720.

National Institutes of Health. 2012. NIH Director's Pioneer Award Program—2013: Frequently Asked Questions. August 7. https://commonfund.nih.gov/pioneer/faq13.aspx#general1.

Nordhaus, William D. 1969. *Inventions, Growth and Welfare: A Theoretical Treatment of Technological Change.* Cambridge, MA: MIT Press.

National Public Radio. 2012. Failure: The F-Word Silicon Valley Loves and Hates. June 19. http://www.npr.org/2012/06/19/155005546/failure-the-f-word-silicon-valley-loves-and-hates.

National Science Foundation (NSF). 2012. *Science and Engineering Indicators 2012*. Arlington, VA: National Science Board.

O'Donoghue, Ted, Suzanne Scotchmer, and Jacques-Francois Thisse. 1998. Patent Breadth, Patent Life, and the Pace of Technological Progress. *Journal of Economics & Management Strategy* 7 (1): 1–32.

Reinganum, Jennifer F. 1989. The Timing of Innovation. In *Handbook of Industrial Organization*, *Volume I*, ed. Richard Schmalensee and Robert D. Willig, 849–908. Amsterdam: North-Holland.

Rockett, Katherine. 2010. Property Rights and Invention. In *Handbook of the Economics of Innovation*, *Volume I*, ed. Bronwyn H. Hall and Nathan Rosenberg, 315–380. Amsterdam and New York: Elsevier.

Romer, Paul M. 2001. Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers? In *Innovation Policy and the Economy*, *Volume 1*, ed. Adam B. Jaffe, Josh Lerner, and Scott Stern, 221–252. Cambridge, MA: MIT Press.

Sampat, Bhaven N. 2006. Patenting and US Academic Research in the 20th Century: The World Before and After Bayh-Dole. *Research Policy* 35 (6): 772–789.

Schumpeter, Joseph. 1942. *Capitalism, Socialism and Democracy*. New York: Harper & Brothers.

Scotchmer, Suzanne. 2004. *Innovation and Incentives*. Cambridge, MA: MIT Press.

Shapiro, Carl. 2001. Navigating the Patent Thicket: Cross-licenses, Patent Pools, and Standard-setting. In *Innovation Policy and the Economy, Volume I*, ed. Adam B. Jaffe, Josh Lerner, and Scott Stern, 119–150. Cambridge, MA: MIT Press.

Shavell, Steven, and Tanguy van Ypersele. 2001. Rewards vs. Intellectual Property Rights. *Journal of Law and Economics* XLIV (October): 525–548.

Stephan, Paula. 2012. *How Economics Shapes Science.* Cambridge, MA: Harvard University Press.

Stern, Scott. 2004. Do Scientists Pay to Be Scientists? *Management Science* 50 (6): 835–853.

UNESCO (United Nations Educational, Scientific, and Cultural Organization) Institute for Statistics. 2012. *Science and Technology Statistical Tables*. http://stats.uis.unesco.org/unesco/ReportFolders/ReportFolders.aspx?IF_ActivePath=P,54&IF_Language=eng.

U.S. Patent and Trademark Office. 2012. Patent Counts by Class by Year: January 1977–December 2011. http://www.uspto.gov/web/offices/ac/ido/oeip/taf/cbcby.htm.

Weyl, E. Glen, and Jean Tirole. 2012. Market Power Screens Willingness-to-Pay. *Quarterly Journal of Economics* 127 (4): 1971–2003.

Williams, Heidi. 2012. Innovation Inducement Prizes: Connecting Research to Policy. *Journal of Policy Analysis and Management* 31 (3): 752–776.

Wright, Brian D. 1983. The Economics of Invention Incentives: Patents, Prizes, and Research Contracts. *American Economic Review* 73 (4): 691–707.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science* 316 (5827): 1036–1039.

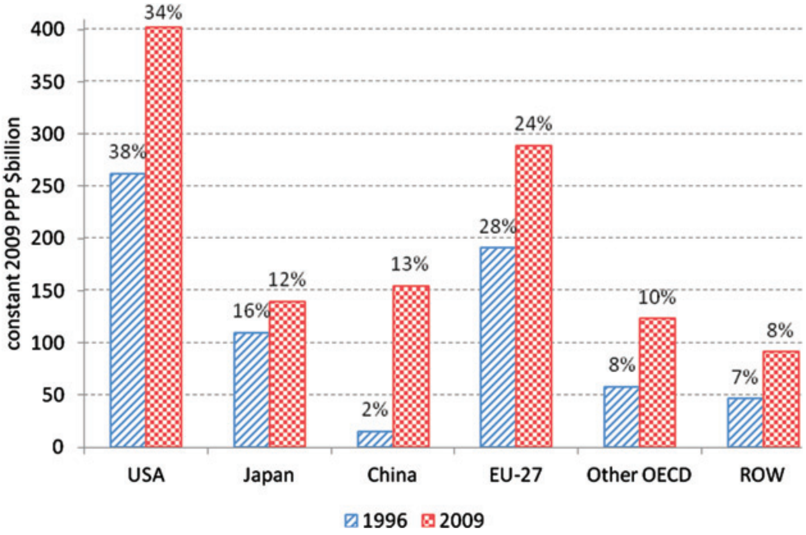
Ziedonis, Rosemarie H. 2008. On the Apparent Failure of Patents: A Response to Bessen and Meurer. *Academy of Management Perspective* 22 (4): 21–29.

Appendix
Research and Development Funds and Innovation Activity: A Brief Overview

The United States shoulders the largest share of global R&D spending, followed by the European Union (EU), China and Japan. Figure A1 illustrates the extent of global R&D by country/region for 1996 and 2009 (the change between these two years has been fairly smooth). World R&D has increased in real terms over this period by about 4.4% per year, with the fastest rate of growth occurring in China and the slowest in Japan. In the United States specifically, figure A2 shows that, whereas total R&D in real terms has been climbing more or less steadily in the last sixty years, there has been a marked shift in the source of R&D investment in the last few decades. In the 1950s and early 1960s, the federal government significantly increased its support for R&D, with its contribution rising from approximately 0.75% to almost 2.0% of GDP, before dropping back to under 1.0% of GDP by 2010. Industry, however, has steadily increased its investments in R&D over the entire period, and overtook federal spending around 1980. The continued rise in industrial R&D has compensated for the decline in federal support, keeping total U.S. R&D levels between 2.5% and 3.0% of GDP.

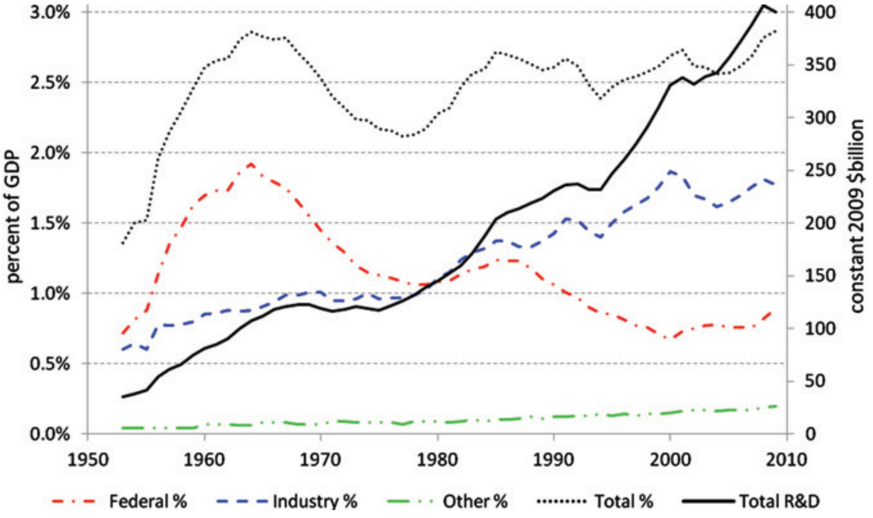
The change in the composition of U.S. R&D funds matters because the source of funds is correlated with the type of R&D performed. This is illustrated in table A1, which shows 2009 R&D expenditure by type of

Figure A1 Total R&D Expenditures (Constant 2009 $ Billion) by Country/Region, 1996 and 2009



Notes: Data are from UNESCO Institute for Statistics (2012). Data labels on columns represent each country/region's percentage contribution to world R&D in the given year. Other OECD includes: Australia, Canada, Iceland, Israel, South Korea, Mexico, New Zealand, Norway, and Turkey. ROW = Rest of the World.

Figure A2 U.S. Total R&D Expenditures (Constant 2009 $ Billion) and Contribution by Sources of Funds (as Percentage of National GDP), 1953–2009



Notes: Data are from the NSF (2012), appendix tables.

research, source of funds, and performing sector. The federal government supports basic research, applied research, and development in roughly equal measure. Universities and other groups (a category that includes nonprofits and state-level government) are also heavily slanted towards

Table A1 U.S. R&D Expenditures ($ millions) by Sources of Funds and Performing Sector, 2009

		Performers				
Source of Funds		**Federal**	**Industry(a)**	**U&C(b)**	**Other Nonprofit(c)**	**Total**
Federal	**Total**	**30,901**	**46,019**	**36,543**	**10,968**	**124,431**
	Basic Res.	*5,507*	*3,890*	*26,050*	*5,004*	*40,451*
	Applied Res.	*8,006*	*9,727*	*7,866*	*4,502*	*30,101*
	Development	*17,389*	*32,403*	*2,628*	*1,462*	*53,882*
Industry	**Total**		**242,820**	**3,279**	**1,258**	**247,357**
	Basic Res.		*13,444*	*2,344*	*698*	*16,486*
	Applied Res.		*33,258*	*767*	*319*	*34,344*
	Development		*196,118*	*168*	*241*	*196,527*
U&C	**Total**			**11,436**		**11,436**
	Basic Res.			*8,173*		*8,173*
	Applied Res.			*2,675*		*2,675*
	Development			*587*		*587*
Other(d)	**Total**			**8,093**	**9,141**	**17,234**
	Basic Res.			*5,785*	*5,075*	*10,860*
	Applied Res.			*1,893*	*2,317*	*4,210*
	Development			*416*	*1,749*	*2,165*
Total	**Total**	**30,901**	**288,839**	**59,351**	**21,367**	**400,458**
	Basic Res.	*5,507*	*17,334*	*42,352*	*10,777*	*75,970*
	Applied Res.	*8,006*	*42,985*	*13,201*	*7,138*	*71,330*
	Development	*17,389*	*228,521*	*3,799*	*3,452*	*253,161*

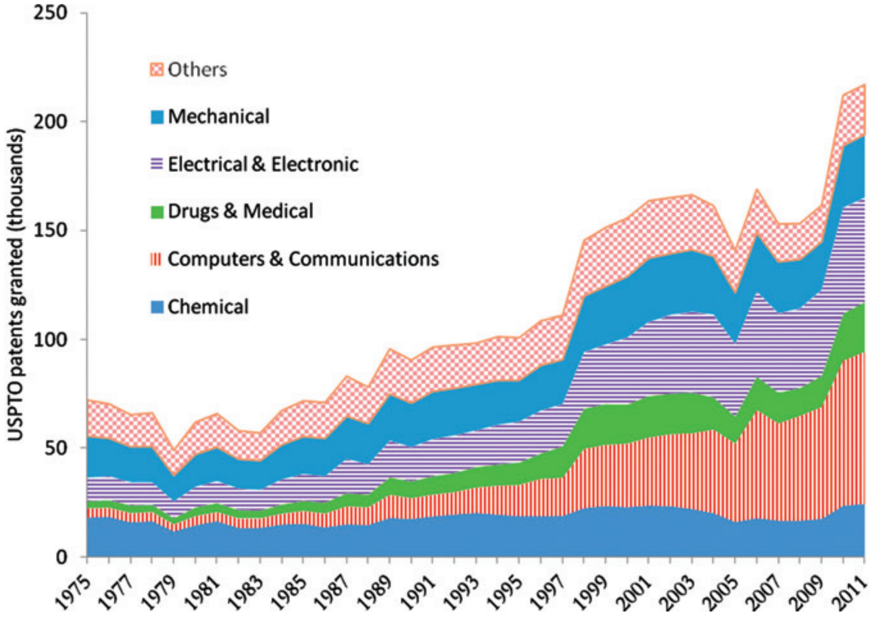
Notes:
(a) Includes industry FFRDCs (federally funded research and development centers).
(b) U&C = Universities and Colleges; includes U&C FFRDCs.
(c) Includes nonprofit FFRDCs.
(d) Includes nonprofit organizations and other government funding.
Source: NSF (2012), appendix tables.

the funding of basic research. Industry, however, is overwhelmingly biased towards financing development.

In addition to its role as the chief source of R&D funds, industry also conducts the lion's share of R&D activities. From the $247 billion that industry spent on research in 2009, $242 billion was kept inside industry, with the remaining $5 billion split between universities and other nonprofit research institutions. Industry also conducted $46 billion of R&D funded by the federal government so that, overall, approximately 72% of total U.S. R&D is conducted by industry. Industry is predominantly engaged in development, which accounts for 79% of its R&D. Basic research remains mostly the purview of universities and other nonprofit institutions, who together conduct 70% of basic research, despite their small share in overall R&D spending (15% for universities, 5% for other nonprofits). Universities and colleges are the second-largest performer of R&D, and are financed from a more diverse set of groups than industry. The majority of funds come from the federal government, followed by internal financing, "other," and then industry. All sources of funds supporting R&D performed at universities primarily target basic research. The third-largest R&D performer is the federal government, which is self-financed, and other nonprofits are the smallest contributors.

The increasing importance of industry R&D is also reflected in the rise of patenting over the last few decades. This is documented in figure A3,

Figure A3 USPTO Patents by Technology Class and Year of Grant, 1975–2011



Note: Data are from the U.S. Patent and Trademark Office (2012) from 1991–2011, and Hall, Jaffe, and Trajtenberg (2001) from 1975–1990.

which reports the number of patents granted by the U.S. Patent and Trademark Office (USPTO) since 1975, and also breaks down such patents by technology class. Patent classes issued by the patent office (of which there are hundreds) are grouped into the six main categories developed by Hall, Jaffe, and Trajtenberg (2001). The annual number of patents granted by the USPTO has risen significantly over the considered period, tripling from approximately 72,000 in 1975 to 216,000 in 2011. It is also notable that, over this period, patenting activity has also shifted dramatically away from an emphasis on "chemical" and "mechanical" patents to "computer and communications" and "electrical and electronic" patents.

Table A2 lists U.S. R&D expenditure by governmental agency and type of research in (fiscal year) 2009. The Department of Defense (DOD) accounts for close to half of all spending, which is massively dedicated to development rather than basic or applied research. The Department of Health and Human Services (HHS), which includes the National Institutes of Health, is the next largest source of funding, at about half the size of the DOD. In contrast to the DOD, the HHS R&D budget is almost entirely devoted to basic and applied research, with the life sciences being the primary area receiving support. The remaining 25% of the federal R&D budget, in descending order of size, goes to: the Department of Energy (DOE), the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA), the Department of Agriculture (USDA), and other federal units. With the exception of DOD and NASA, these agencies devote the majority of their budget to basic and applied research. As for the allocation of the latter funds, HHS and USDA heavily support the life sciences, whereas other agencies split their support more evenly across scientific domains.

Table A2 U.S. Federal Government R&D Obligations ($ Millions) by Selected Agency, Character of Work, and S&E Field, FY2009

	DOD	HHS	DOE	NSF	NASA	USDA	Other	*Total*
Development	61,307	94	2,702	0	4,234	192	1,112	**69,640**
R&D Plants	117	152	1,672	830	0	775	799	**4,345**
Basic & Applied Research of Which:	6,806	35,490	7,188	6,095	1,703	2,078	4,092	**63,453**
Environmental science	390	545	373	1,079	395	19	949	**3,751**
Life science	924	27,930	349	934	97	1,737	1,298	**33,268**
Mathematics and computer science	913	187	1,002	1,253	29	10	214	**3,607**
Physical science	833	297	2,612	1,181	469	106	322	**5,820**
Psychology	69	1,904	0	11	6	0	96	**2,087**
Social science	39	294	0	253	0	160	300	**1,046**
Other science	145	2,579	375	366	87	0	91	**3,644**
Engineering	3,493	1,754	2,476	1,019	620	46	822	**10,230**
Total R&D	**68,230**	**35,736**	**11,562**	**6,925**	**5,937**	**3,045**	**6,003**	**137,438**

Notes: DOD = Department of Defense; HHS = Department of Health and Human Services; DOE = Department of Energy; NASA = National Aeronautics and Space Administration; USDA = Department of Agriculture.
Source: NSF (2012), appendix tables.

In addition to the role of patents and contracted research, this paper also examined the role of prizes as a form of innovation incentive. One must note, however, that prizes account for a negligible percentage of R&D spending at this time. The 2010 America COMPETES Reauthorization Act did establish a system for federal agencies to offer innovation inducement prizes, and with time we may perhaps see prizes forming a larger share of R&D expenditure. However, as of November 17, 2012, there were 227 prize challenges offered on the federal prize clearinghouse website, www.challenge.gov, valued at under $50 million in total, or under 0.03% of total federal R&D outlays.