How Disclosure Policies Impact Search in Open Innovation

Kevin J. Boudreau (London Business School)

Karim R. Lakhani (Harvard Business School)

Abstract

Most of society's innovation systems–academic science, the patent system, open source, etc.–are "open" in the sense they are designed to facilitate knowledge disclosures amongst innovators. An essential difference across innovation systems, however, is whether disclosures take place only after final innovations are completed or whether disclosures relate to intermediate solutions and advances. We present experimental evidence showing that implementing intermediate versus final disclosures does not just create quantitative tradeoffs in shaping the rate of innovation. Rather, it qualitatively transforms the very nature of the innovation search process. Intermediate disclosures have the advantage of efficiently steering development towards improving existing solutions, but curtails experimentation and wider search. We discuss comparative advantages of systems implementing intermediate versus final disclosures.

1 Introduction

Innovation has been described as an on-going, cumulative process of discovery across an endless knowledge frontier–in which one innovator often builds upon the progress of others. The range of society's institutions for encouraging and governing innovation–the patent system, academic research, the system of government research labs, open source projects, public contests, etc.–are each designed to both create incentives for innovators to exert effort and to ensure that the discoveries are eventually disclosed and available for downstream innovators (Scotchmer 2004). While past research on the role of disclosure in innovation has focussed on the effects of breadth and timing (e.g., Mazzoleni and Nelson, 1998; Lerner, 2000; Varian, 2005) and specific means of disclosures in particular innovation systems (e.g., Kitch, 1977; Dasgupta and David, 1994; Lerner and Tirole 2005), in this paper we wish to highlight and study how disclosure policies can fundamentally distinguish altogether different innovation systems. This is the distinction between disclosures of "final" or working innovations and disclosures of "intermediate" advances, generated *before* the final innovation may be complete.

Our traditional innovation institutions tend to be geared towards facilitating final disclosures in some completed, final and working form, as in a working invention, a completed and vetted academic paper, an enabling technology platform or biological organism. However, recently there have been more conspicuous examples of systems geared to implementing disclosures of intermediate advances–early working solutions, instructions, materials, research inputs, etc.. For example, academic institutions participating in the Human Genome Project agreed to the "Bermuda Principles," committing themselves to each release all new sequence data to a public database within 24 hours of discovery (Contreras, 2011), significantly well ahead of any peer reviewed publication. Also well known and noted is the tremendous mobilisation of shared knowledge in open source software development projects. Formal and informal frameworks governing open source software development projects, such as the GNU General Public License, not only grant freedoms to use, study, share, and modify code, but also obligate subsequent software modifications to be made available for others (O'Mahony, 2003; von Hippel and von Krogh, 2003; West, 2003).[1]As the examples illustrate, intermediate disclosures first discretely differ from final disclosures in their timing. Final disclosures are bound, by definition, to occur *after* the final preparation of a completed work–at their earliest (and typically considerably later). Intermediate disclosures can occur as early as they are generated, sometimes even virtually instantaneously. Further, whereas final disclosures typically involve some standard, integral form of knowledge transfer, intermediate disclosures might also allow for greater breadth of disclosure or even smaller quanta of knowledge, as in partial results and progress, negative results, methods, data, and so on.

There are now growing calls for our academic and industrial science systems to move further towards intermediate disclosures in hopes that these might increase productivity and innovativeness (Nelkin, 1982; Lin, 2012; Nielsen, 2012; Royal Society, 2012). However, we yet have limited systematic research to indicate under what conditions this might be true. In this paper, we attempt to address this gap by presenting controlled experimental evidence on how intermediate versus final disclosure policies affect innovation outcomes across populations of innovators working under either regime.We begin by demonstrating that an intermediate disclosure policy, where all or some of one's work is available for reuse by others before the final innovation is complete, lowers incentives, discourages effort and reduces participation, while enhancing knowledge transfers–revealing a classical incentives-versus-reuse tradeoff produced by varying disclosure policies (e.g., Scotchmer 1991; Bessen and Maskin, 2009).

Our main contribution, is then to demonstrate an intermediate disclosure policy initiates a qualitatively different sort of innovative search process than does a policy of final disclosures. On the one hand, a policy of final disclosures (i.e., no intermediate disclosures)

[1]Within social sciences, the FIVE project, organised by professors Constance Helfat and Steven Klepper, was organised with the goal of promoting research on firm and industry evolution by making available data sets (see, for example, Murmann 2007; Thompson 2012). Many social science journals now ask for data sets to be disclosed. In other examples, Wikipedia and other public and corporate wiki's implement a system whereby all partial changes are instantaneously reflected in the presentation of the wiki. Within arts and culture, we see analogous changes with growing governance of books and literature under Creative Commons licenses. Within popular music, it is a growing norm to make available elemental components of finished works (individual music and vocal tracks) and instructions (lyrics, music transcriptions and tablature).

allows innovators to make progress with a greater degree of independence (and ignorance) of others' development choices–or "parallel experimentation" (Abernathy and Rosenbloom, 1969; Nelson, 1961; Terwiesch and Xu, 2008, Boudreau et al., 2011; Scherer, 2011). This independence, in addition to high incentives implied by the incentives-versus-reuse tradeoff, mean we should expect higher levels of experimentation and wider breadth of search under final disclosures. On the other hand, a policy of intermediate disclosures accelerates knowledge sharing amongst the population of innovators. Intermediate disclosures also provide signals of the most productive existing technical approaches, generating a greater degree of common knowledge regarding the shape and outlines of the technological frontier. Hence beyond just enabling reuse, the signals and information created by intermediate disclosures can lead to "coordinated" patterns of innovation and experimentation–and tend the population of innovators working under such a system to converge on established technical paths.

Given that intermediate versus final disclosure policies tend to be most important across altogether different systems (e.g., open source versus academic science, etc.), there are inherent challenges to making all-else-being equal comparisons in naturally occurring contexts in the economy. Therefore, we implement our test in a field experiment to isolate relevant differences and resulting outcomes. In our experiment, a total of 733 subjects–comprised of mathematicians, software developers, scientists, and data scientists–were tasked with developing solutions for a challenging bioinformatics algorithmic problem, from researchers at the Harvard Medical School, over two weeks. Subjects were randomly assigned to independent controlled groups (controlling the composition of the group of prospective entrants, problem type, time, incentives, and institutional features beyond disclosure policy). Separate experimental groups differed in terms of whether they worked under a policy of disclosures of intermediate solutions (i.e., making available their intermediate working solutions) or a policy under which intermediate solutions were kept secret and only final solutions were disclosed at the end of the experiment. Those developing the best solutions within each independent group received cash prizes and public acclaim. The contest-style payoff structure was intended to reflect the typically skewed rewards for those providing best innovation solutions (e.g., Scherer and Harhoff 2000). Our research design also relies on implementing a series of appropriate measures for plainly revealing relevant differences. Following this approach, our analysis is able to rely upon simple and explicit comparisons of patterns to reveal readily apparent differences that emerge across populations subject to different disclosure policies. In this sense, our analysis here does not rely on subtle econometric manipulations or modeling in order to demonstrate conclusions.

In our findings, we first document both sides of the incentives-versus-reuse tradeoff. We then present a range of aggregate population-level patterns and micro-level evidence to reveal the striking differences in the search process under the different disclosure policies. We discuss how these differences lead to different levels of productivity and innovativeness than what would be predicted by the incentives-versus-reuse tradeoff on its own. Our discussion elaborates on how these results provide further clues on the "division of labour" among different kinds of innovation systems in addressing society's (and individual organisations') innovation problems.

The results contribute to a wide range of work related to the design of open and distributed innovation systems, knowledge-sharing and processes of on-going innovation (Chesbrough 2003; von Hippel 2005; O'Mahony 2003; Lakhani and von Hippel 2003; von Hippel and von Krogh 2005; Murray and Stern, 2007; Rysman and Simcoe, 2008; Boudreau 2010; Furman and Stern, 2011; Dasgupta and David, 1994; Williams, 2013). However, rather than elaborate on a particular institutional context, here we examine an essential difference in disclosure policies across innovation systems. This work also adds to abundant past research on questions of timing and breadth of disclosures (as in the duration of patents and copyright) (Mazzoleni and Nelson, 1998; Lerner, 2006; Varian, 2005); however, here we distinctly examine case where the timing is *before* the creation of a final innovation. Our work also closely relates to a stream of papers that have documented effects of granting greater disclosures and access to upstream knowledge (Murray and Stern, 2007; Furman and Stern, 2011; Williams, 2013). However, here we document a wider range of tradeoffs under precisely controlled conditions. We also are able to clarify reasons for caution in interpretting results as these. We also contribute micro-level evidence on mechanisms underlying cumulative innovation and parallel experimentation, empirically demonstrating the usefulness of characterisations of innovation as "search" (Nelson, 1969, 1971; Abernathy and Utterback, 1978; Sahal, 1981; Dosi, 1982) in the data. We substantiate intuitive tradeoffs between policies that stimulate cumulative innovation (e.g. Osterloh and Rota, 2007; Murray and O'Mahony, 2007) versus parallel experimentation (e.g. Nelson, 1961; Abernathy and Rosenbloom, 1969; Scherer 2011), while clarifying a range of mechanisms (incentive effects, information and signaling, and knowledge transfers) underlying them. Secondarily, we provide explicit micro-level evidence of uncertain search processes, whereas prior literature has tended to focus on theoretical models and simulations (e.g., Levinthal, 1997).

2 Related Literature & Predictions

Questions of disclosure, knowledge transmissions and spillovers among innovators and innovation outcomes have been addressed by a number of distinct literatures–mostly compartmentalised by distinct institutional approaches to innovation, i.e. science, patents, user innovation, platforms, and so on. For example, as regards academic science, the research examines the rich institutional design of the "Republic of Science" to explain how this system both creates incentives and implements disclosures (e.g., Dasgupta and David, 1994; Stephan 1996). A long tradition of research has studied how institutions undergirding industrial competition shape innovation outcomes, as in questions of the timing and breadth of patent protections (e.g., Nordhaus, 1972; Klemperer, 1990) or copyright (e.g., Lessig, 1999; Klein et al., 2002) or informal protections (e.g..: Cohen et al., 2000). Implicit in many of our theories of market competition and innovation is the assumption of a good deal of involuntary "leakage" and spillovers of productive knowledge (e.g., Caballero and Jaffe, 1993). Distinct bodies of literature have begun to explore and document the functioning and governance of open source software projects (e.g., von Hippel and von Krogh, 2003; Lerner and Schankerman, 2010). Relative to these analyses focusing on particular institutions, our paper instead attempts to make broad comparisons across different innovation systems.

Apart from the question of an innovation system imposing a disclosure policy (i.e., rules to abide by), research on "open innovation" also documents instances of discretionary, elective and strategic disclosures made by innovators of their own volition. For example, Chesbrough (2003) highlights "open innovation" as purposive knowledge flows into and out of the firm through a variety of means, including licensing. Keld and Salter (2006) continue in this tradition, documenting exploitation of a variety of knowledge sources outside the firm. Haeussler et al. (2009) examine disclosures made by scientific researchers to close confidants and associates, as well as to broad audiences prior to publication. A number of studies document informal "collective invention" groups of inventors, wherein inventors have often elect to share ideas whilst promoting innovation in an industry (Allen, 1982; Osterloh and Rota, 2007; Fauchart and von Hippel, 2008; Meyer, 2013).

The Classical Incentives-versus-Reuse Tradeoff The deep specialisation's of each of these above mentioned literatures and the evolved complexity of the institutions studied creates certain challenges in attempting to make comparisons across systems. Nonetheless, it is fair to point out now well accepted first-order tradeoffs that have been identified across these literatures.

The classical way of thinking about the effects of varying levels of disclosures and sharing is as a tension between ex-ante[2] incentive provision and ex-post[3] reuse–or an *incentives-versus-reuse tradeoff* (see Scotchmer 2004, ch. 5). Of course, where knowledge disclosures are implemented (and access granted to disclosed knowledge, too), innovators may borrow, reuse and build upon each others' ideas, technologies and methods (e.g., Rosenberg, 1984; Mokyr,

[2] I.e., prior to the discovery of a productive advance.
[3] I.e., after the realisation of a productive advance.

2002; Murray and O'Mahony, 2007). However, because greater disclosures can degrade rewards, recognition and payoffs to upstream innovators, there may be a cost of instilling lower incentives for innovators to make risky investments in the first place (e.g., Green and Scotchmer, 1995; Bessen and Maskin, 2009).

To date, empirical papers related to incentives-versus-reuse have tended to focus on just one side of the tradeoff at a time. For example, there has been considerable interest in evaluating the extent to which patents, copyrights, and other protections (or, conversely, disclosures) have affected the level of inventors' incentives and innovative output (e.g., Cohen et al. 2002; Moser 2005; and many more). Also reflecting the empirical importance of incentives, we see in cases of high disclosures a tendency to implement "compensating mechanisms" to offset what would otherwise be faultering incentives. For example, academic science, open source software development and wikis each implement attribution systems. By having contributors "sign" their work, it is possible to collect benefits in the form of reputation, especially when others reuse one's work (Lerner and Tirole, 2005; Dasgupta and David, 1994; Stephan, 1996). Also, in contexts with high disclosures, innovation systems appear to also often rely on attracting individuals who are motivated on grounds other than capturing traditional rewards and who are susceptible to being socialized in an ethos of sharing with peers (e.g., Roberts et al., 2006; O'Mahony and Murray, 2007; von Krogh et al., 2012; Meyer 2013).

Mostly separate strands of literature document the "reuse" half of the incentives-versus-reuse tradeoff. For example, von Hippel and colleagues (e.g., von Hippel 1988, 2005; Lakhani and von Hippel, 2003; Baldwin et al., 2007; von Hippel et al., 2013) have documented the existence and functioning of groups of "user innovators" who commonly disclose innovations made by building on and modifying existing products. This is somewhat analogous to research on technology platforms, which focuses on the reuse of some upstream technological set of assets by downstream innovators who build upon the platform to commercialise a range of products. Boudreau (2010), for example, documents how increased opening and access to mobile computing and smartphone platforms through a variety of contractual and technical approaches multiplied rates of product development by several times. Rysman and Simcoe (2008) document outsized citations to technologies that become designated as platform technologies. Related to a wider range of technology categories, Galasso and Schankerman (2013) find that technologies coming off patent as a result of legal challenges experience a surge in reuse thereafter.

As regards knowledge reuse in scientific research, a number of studies have now looked at before-and-after effects of granting more straightforward disclosure and access to upstream research. Murray and Stern (2007), for example, show a significant decline in citations for papers that have their underlying discoveries and knowledge under patent protection. Similarly, Williams (2013) finds that gene sequences once subject to private firm Celera's intellectual property experience persistently lower follow-on publications and rates of follow-on patenting that cite those sequences–even once Celera's restrictions are no longer maintained. Huang and Murray (2009) find evidence suggesting the depressing effect of patents on on-going research publication rates on specific genetic sequences is exacerbated in cases with broader patent scope, private sector ownership, patent thickets, and fragmented patent ownership. Furman and Stern (2011) find that submitting biological resources materials to biological resource centers (BRC) resulted in a 50-125% boost in citations to publications associated with those materials. Hence the literature finds, as one might expect, that that greater disclosures and rights of access leads to significantly greater reuse, across a range of technological and institutional contexts.

Inasmuch as intermediate disclosures tend to produce a greater degree of disclosures (i.e., wider breadth and/or earlier timing) and, at the same time, tend to degrade rewards to original upstream innovators, we should expect differences between intermediate and final disclosures to be subject to the classical incentives-versus-reuse tradeoff.

> *Prediction 1:* Implementing an intermediate disclosure policy sets into a motion an incentives-versus-reuse tradeoff, lowering incentives to participate and exert

effort while at the same time increasing the reuse of innovative knowledge across innovators.

Disclosure Policies and the Uncertain "Search" Process The earlier discussion about the tradeoff between levels of incentives versus levels of reuse have implied in them a characterisation of the innovation process as one of trading these levels as "inputs" in some risky innovation "production function". However a parallel literature has characterised the innovation process as a "search" for the best solution(s) within an uncertain solution space. Scholars in this tradition, rather than just drawing attention to *levels* of effort and knowledge reuse, have focussed questions on the *direction* and performance of technical approaches that are being pursued by innovators. The literature in particular has drawn attention to the overall patterns and distribution of experimentation and choices of "pathways", "trajectories," "paradigms" pursued by *populations* of innovators (cf. Kuhn, 1962; Abernathy and Utterback, 1978; Sahal; 1981, Dosi, 1982). No theory or model of individual or 'average' behaviour quite addresses this question.

This view of innovation as search figures prominently in a wide range of research on technical innovation and scientific discovery since the inception of modern scholarship on the topic (Kuhn, 1962; Nelson, 1971; Abernathy and Utterback, 1978; Sahal, 1981; Dosi, 1982; March, 1991; Levinthal, 1997; Teece, 2008; von Tunzelmann et al., 2008; Erat and Krishnan 2011). The literature actively considering these questions demonstrates stark differences in outcome that turn on the direction of technical search. For example, Dosi (1982), invoking Kuhn (1962), articulated that innovation inside of firms occurs within well-accepted technological paradigms that determine a trajectory of advancement and progress; with the R&D actions of many firms circumscribed by the paradigm that they inhabit and which "focusses the eyes and efforts of technologists and engineers in defined directions" (Dosi, 1982, p. 158). Dosi further noted that within the economy there may exist simultaneous multiple technological paradigms attempting to solve the same underlying innovation problem by providing the example of the possibility of both semiconductor technology and thermoionic valve technology as competing approaches towards advances in computation (Dosi,1982, p. 159). Other examples of the presence of various paradigms in technological development have been shown for a variety of technological systems including aircraft (Vincenti, 1994; Murman and Tushman, 1998) and bicycles (Bijker and Pinch, 1987). Taking this perspective to the level of market and industry competition, Utterback and colleagues' (Utterback and Abernathy, 1975; Utterback and Suarez, 1993; Utterback, 1994) characterisation of the period of technological ferment in many industries, before the emergence of a dominant design, is akin to a population-level search for the best solution in light of many potential technical paths being explored.[4]

Formalizations of the search metaphor for both innovation-related problem solving and generalised managerial decision making began with seminal works such as those of Simon (1962) and Newell and Simon (1972) and has taken fruition in formalised models within a distinct literature on organisational search (e.g., Levinthal, 1997; Rivkin and Siggelkow, 2003, 2007;). Using Kaufman's (1993) "N-K" modelling apparatus, researchers have "mapped" problem-solving and innovation in terms of consisting of a large set of problem-solving decisions, with some degree of interactions among them, that together produce a complex solution (payoff) space (Fleming and Sorensen, 2004; Nickerson and Zenger, 2004; Sorensen et al., 2006). Innovators then search this landscape for the best performing solutions, by trying different combinations of technologies and decisions (c.f.: Fleming and Sorensen, 2004) and depending on the existing stock of knowledge they possess, i.e., where they start on the landscape, and their search heuristics. As a result, individual innovators may find themselves in local optima or maximally performing global peaks (Rosenkopf and Nerkar, 2001).

Here, we consider the implications of this richer characterisation of the innovation process as more than just a production process, but as one of a search process in many possible

[4]Recent such examples include the choice between plasma, liquid crystal display and light emitting diode technologies for digital television and the various alternative ways to achieve more environmentally sustainable automotive engines through the development of hydrogen gas, gas-electric hybrids and all electric vehicles.

"directions" within an uncertain solution space. This raises the question of the impact of disclosure policies on the search behaviour of individuals and the patterns of search of broader population of innovators that are subject to a given disclosure regime. A disclosure policy involving only disclosures of final works, where there is no sharing of intermediate solutions or knowledge, should result in more independent and therefore less correlated search approaches taken among individual innovators. Each innovator, working with their own private and idiosyncratic knowledge and abilities, will independently search the solution landscape and depending on their search heuristic either get trapped within a local optima or discover a high-value solution. In essence, at the population level, with each innovator is in this case trying to solve the same problem; a "parallel" search process will be undertaken with the potential for a wide distribution of outcomes being explored and a great many combinations of technologies being investigated (Abernathy and Rosenbloom, 1969, Nelson, 1961; Terwiesch and Xu, 2008, Scherer 2011; Boudreau et al., 2011).

> *Prediction 2*: In a final disclosure regime, the population of innovators will engage in patterns of relatively broad, independent experimentation, resulting in a wide range of performance outcomes and technical approaches.

In contrast, a policy of intermediate disclosures spills over productive knowledge to the wider population of innovators. More than this, intermediate disclosures also telegraph the direction and outcomes of others' experimentation, creating signals about the outlines of the technical frontier, the activity and "positions" of others along that frontier and the outcomes of experimentation. Therefore, rather than producing greater independence, intermediate disclosures should lead to greater coordination and non-random *patterning of choices* of technical approaches. In principle, coordinated and informed strategic choices could lead to more deliberately differentiated experiments across the solution space (cf. Salop 1979, Hoteling 1929). However, the history of scholarship in this area suggests otherwise. As an empirical matter, we tend to see that technical solutions, trajectories, paradigms or pathways of innovators tend to converge on similar approaches, once viable and productive approach has been identified–allowing continued investments to proceed under palpably lowered uncertainty (e.g., Kuhn, 1962; Dosi, 1982; Utterback and Abernathy, 1972). Therefore, we anticipate that intermediate disclosures should produce greater "convergent" accumulation of technical development and limit the breadth of search undertaken:

> *Prediction 3:* In an intermediate disclosure regime, innovators will engage in limited experimentation and instead tend to converge on established well-working technical approaches.

3 Experimental Design and Methods

Rather than just a source of incremental "tweaks" and tuning of disclosures, the distinction between intermediate and final disclosures is most salient *across* altogether different innovation systems. This creates inherently difficult research design challenges. For example, attempting to compare the (intermediate) disclosures of GPL-governed open source software projects in relation to (final) disclosures of working papers in academic research must recognise concurrent differences in technical domains, the 'basicness' and kinds of problems addressed, the nature and size of the pools of innovators each system draws on, the basic interests and economic motivations of actors within each system, norms and culture shaping behaviours, the particulars of the incentive system put into place and so on. Correspondingly, any innovation or governance system should be understood as a complex mesh of incentives, constraints, rules, culture, norms, information and supporting facilities that should work together to produce desired outcomes (see, for example, Dasgupta and David 1994).

Therefore, it might be all but impossible to make *ceteris paribus* comparisons of intermediate versus final disclosure policies across altogether different institutions or innovation regimes. At the same time, it is precisely because these particular sorts of differences in disclosure policies are most salient across systems that we might suspect they are worthy of careful study and could potentially explain their comparative strengths. With these points in mind, we pursue an experimental design. We also necessarily abstract from a range of varying institutional details in order to make relevant *ceteris paribus* comparisons, setting each assignment group within contest-style payoffs to reflect skewed rewards to best-performing innovators. We do not attempt to explore endless variants of ways and degrees of implementing intermediate disclosures, but rather test key predictions implementing regimes either with or without intermediate disclosures. Here we describe details of the design.

3.1 The Problem Addressed by Subjects in the Experiment

The innovation challenge to be addressed in our field experiment involved designing a bioinformatics algorithm for comparing and annotating a very large series of genomic sequences –a problem involving the processing of large amounts of data, minimising errors, dealing with possible errors in data, solving within constrained computational resources and minimising the amount of time (see Lakhani et al. 2103). The problem itself was sourced from researchers at Harvard Medical School and was part of a systematic investigation into the use of open innovation approaches within the broader academic research enterprise (Guinan et al., 2013). Subjects developed *de novo* solutions in computer code. They faced the constraint that their code be able to annotate 10^5 genetic sequences on a desktop computer. Conditional on meeting this constraint, the quality of solutions could then be judged on accuracy and speed. The experiment and time to develop solutions was two weeks.

A number of features of this problem make it appropriate for our purposes, and particularly for studying processes of innovation, where both continuous advance and experimentation might play a role. As a problem sitting at the intersection of software development, mathematics, computer science, and biology, it is nontrivial and challenging. It is a sort of optimisation problem whose solutions should attempt to eke out incremental gains–rather than a final analytical solution that is either correct or incorrect. It is therefore a complex, data-intensive numerical optimisation problem that cuts across knowledge domains, and one encountered in a wide range of contexts, including both industrial and academic applications. The problem has itself been addressed in a cumulative innovation process within the academic literature, beginning a couple of decades ago as gene sequencing got underway (Altschul et al. 1990).

More generally, the choice to focus on algorithm development allows us to treat intermediate solutions themselves as the primary input to subsequent development, within a trial-and-error and learning process. This is true whether intermediate solutions are drawn from others' or whether subjects learn from their own trial and error experimentation. (Conducting an analogous experiment in which intermediate work took the form of physical or living materials, such as research mice, would make carrying out the experiment far less practicable for obvious reasons.) Working in digital format also allows solutions to be codified and recorded in computer instructions, and scored in an automated system, as will be described.

3.2 The Experimental Context

We ran the experiment on an online platform in order to implement alternate disclosure policies and otherwise controlled institutional contexts, to make precise assignments to allow us to match the pools of prospective participants in each treatment group, to allow fine-grained micro-measures at the level of individual participants, and to allow us to recover and observe each intermediate and final solution. TopCoder.com, the online platform site, provided a particularly attractive context because it regularly runs algorithmic optimisation contests for corporate, academic and government clients and can draw from a wide range of algorithmic developers, data scientists, software developers and mathematicians from a range of industries and academic contexts. We recruited subjects through the TopCoder website and distribution lists, explaining the duration of the experiment, payoffs, and that the problem would be an algorithmic optimisation problem developed in conjunction with the Harvard Medical School (without stating the nature of the problem). Among the 733 individuals who participated in this experiment, roughly half (44%) were computer and data science professionals and the remaining were students at various levels of achievement. Participants came from 69 countries.

A number of features of the environment were held constant across assignment groups. All development and interactions took place on the online platform. Participants were given the problem statement at the start of the experiment, along with a description of how solution accuracy and speed would translate to quantitive solutions scores. However, subjects received more direct feedback on their solution code simply by submitting and directly observing how their algorithm designs translated to solution scores when assessed against an automated test suite on the platform. It was not possible to receive direct feedback on the quality of the submission "off line." Subjects could compile and test their code in a wide range of leading computer languages. Subjects' final submissions were taken as the basis for determining winners. They could submit as many intermediate solutions as they pleased, as a means of receiving feedback in their trial-and-error development.

3.3 Disclosure of Intermediate Solutions versus Final Solutions

The key idea of the experiment is to compare how similar independent pools of subjects (prospective participants in an innovation process) are exposed to the opportunity to work on a given problem, and to work under similar institutional conditions, at the same time–but under differing disclosure policies.

Under Intermediate Disclosures, the system allowed subjects to refer to a catalogue of existing solutions within the same experimental group, as solutions were submitted and scored. This could be done simply through the same web interface on which they conducted their development. The list or catalogue identified solutions by their score, submitter and time of submission. Clicking on this catalogue immediately directly reveals the algorithmic code. Thus, we implemented a relatively simple and frictionless system of disclosures, in order to better ensure that disclosures and knowledge reuse would indeed figure prominently in dynamics. (We also wished to minimise the possibility that frictions, coordination costs or transaction costs could play a role in this design, as those are not the mechanisms under study here.)

Rather than completely abstracting away institutional features of systems featuring intermediate disclosures, we chose to implement an attribution system in Intermediate Disclosures. The attribution system should not be understood as systematically altering results, but rather to serve as a minimum "compensating mechanism" (as practiced in naturally occurring systems like academic citations and patents) to limit the possibility of a severe drop in incentives created under an Intermediate Disclosures policy. Subjects reusing ideas were instructed to "cite" the solutions they studied, learned from or built upon. Ex-post, we observe that all winning solutions cited other solutions and this directly mapped to the solutions they indeed studied at. Under the No Intermediate Disclosures, winning solutions were simply posted after the experiment. During the experiment, the platform provided no technical support or catalogue for sharing solutions. To minimise any off-platform sharing, subjects were explicitly instructed not to share solutions under threat of being disqualified.

For both kinds of disclosure regimes, both payoffs were awarded after each week. Rewards were both monetary and reputational. As regards cash prizes, the top five positions were allocated a total of \$1000 in cash prizes (\$500, \$250, \$125, \$75, and \$50). In the case of the regime with intermediate disclosures, half of these prizes were allocated to those cited by the winners. Within the regime without intermediate disclosures, winners kept 100% of each prize listed here. Regarding reputation rewards, winners and those cited were also publicly announced on the TopCoder website. (It is common for such accomplishments on the TopCoder platform to be listed on winners' c.v.'s.).

3.4 Assignment of Subjects

Above all, we prioritised creating assignment groups that were as large as possible. This is because the goal in this experiment is essentially one of simulating entire innovation systems–in which entire populations of prospective or possible innovators are confronted with different regimes in which they might (or might not) choose to innovate. Depending on the regime, they then make a choice to participate and how to do so. To maximise the size of assignment groups we created just one large "Intermediate Disclosures" regime and just one "No Intermediate Disclosures" regime. But given that choosing maximally large groups means trading off against replication, our strategy was to just create one other supplementary regime simply to provide a means of assuring results found in these primary comparison groups were not somehow eccentric. Therefore, we constructed a "Mixed" regime, which had no intermediate disclosures in the first week, but implemented intermediate disclosures in the second week of the experiment. We focus on reporting results in our primary comparison regimes, but later report results from the Mixed regime to reveal the regularity of results. Assignment of the 733 individuals across the three regimes was done with the interest of minimising any differences, first matching on TopCoder skill ratings and randomising on unobserved characteristics. There were 245 subjects assigned to Intermediate Disclosures, 244 to No Intermediate Disclosures and 244 to Mixed.

4 Data

The most demanding measurement challenge to address in tackling our research question regarding the nature of the innovation search process is to observe both the quality and the technical approaches employed by each subject. As regards solution quality, we simply report the quantitative score assigned to each submission based on the automated platform test suite (a summary value conveying an amalgam of both speed and precision). To develop a meaningful indication of technical approaches employed in different solutions, we hired three Ph.D.-level experts to examine each of the 654 intermediate and final solutions (submitted by 124 out of the 733 subjects). These experts identified ten key elemental computational techniques used within the population of solutions (see Table 1). Each submission was coded in terms of a 10-digit binary code representing a combination of techniques (see Lakhani, et al. 2013). There were 56 unique combinations developed across the entire experiment.

<Table 1>

At the level of individual subjects, we observe whether they choose to enter and actively participate, proxied by whether they submitted at least one solution. Conditional on entering, we observe levels of effort both in terms of the number of submissions and also a self-reported measure of hours worked (based on a questionnaire administered after the contest and prior to revealing results. We also observe the skill rating received by participants on the TopCoder platform prior to the experiment, based on an Elo-based system (Maas and Wagenmakers, 2005), that estimates skill on the basis of historical performance in similar algorithmic problem-solving exercises.[5]

5 Analysis

5.1 Baseline Patterns: The Incentives-versus-Reuse Tradeoff

In this subsection, we begin by quickly presenting evidence to summarize the existence of an incentives-versus-reuse tradeoff, before proceeding in the following subsection to begin to examine the evidence in relation to our main question related to the process of search. We straightforwardly find evidence consistent with Prediction 1 from Section 2.

Intermediate Disclosures led to clearly lower levels of participation and activity, consistent with a drop in incentives. In Intermediate disclosures, the rate of active participation to was 26% lower that than in Final Disclosures (14% instead of 19%, significant at $p = 10$; or, 32 versus 46 subjects in absolute numbers). Our measure of effort, the number of submissions per participant was 56% lower (3.9 instead of 6.9 solution submissions, significant at $p = 1\%$),

[5]The average participant engaged in dozens of problems prior to the experiment. The Elo system is standard in a range of contexts from chess grandmaster tournaments to US College Bowl systems to the National Scrabble Association and the European Go Federation.

and the number of self-reported hours worked was 29% lower (10.0 hours instead of 14.1 hours, significant at $p = 1\%$, based on a 60% response rate of participants).[6]

On the other side of the incentives-versus-reuse tradeoff, Intermediate Disclosures led to greater reuse–by design. Under Intermediate Disclosures there were a great many knowledge transfers of intermediate solutions using the relatively frictionless ("click-through") design of the system. In all, 30 of the 32 active participants clicked-through to peak at others' solutions, peaking at a total of 1024 solutions.[7] The bulk of peaking concentrated on early submissions and highest scoring solutions, with a general decline in peaking levels over the duration of the experiment. Four-fifths of peaking occurred in the first half of the experiment. Of course, under Final Disclosures, there were no intermediate disclosures at all.

Also consistent with a tradeoff between incentives-versus reuse, whereas the tradeoff implies no general claims about whether any particular form of openness or secrecy should generally produce more innovation (i.e., which side of the tradeoff should dominate), the theory implies there should be disproportionately high problem-solving performance per effort under Intermediate Disclosures. Consistent with this implication, the average score in Intermediate Disclosures is indeed 1.36 points higher (significant at $p = 1\%$) when controlling for the level of effort and skill rating of the submitter. In fact, even without controlling for skill and effort, both the average and maximum score are higher under Intermediate Disclosures.

5.2 Main Results: Differences in the Innovation Search Process

In following subsections, we proceed to the paper's main prediction: that there are qualitatively distinct kinds of search processes initiated under the different disclosure policies, above and beyond tradeoffs between incentives-and-reuse. Under Intermediate Disclosures, we predicted more cumulative learning and convergence or approaches; under No Intermediate

[6]All significance results reported in this paper are based on robust standard error estimates.

[7]A large number of subjects who did not eventually enter an participate (44) also peeked at solutions, indicating other possible motives or informational value of peeking.

Disclosures, we predicted a greater level of parallel experimentation. Given our prediction here is of qualitatively distinct patterns that should be easily recognisable, our approach here is first to simply present descriptive patterns at the aggregate population level. To the extent the hypothesised patterns are indeed present and important, we should require little or no eonometric manipulation to reveal differences. In following subsection, we examine the behaviours of individuals in greater detail.[8]

To appreciate differences in search patterns, we first examine trajectories of performance. Figure 1 reports all submissions made within both of our main comparison regimes using dots. Dots are joined with lines to show submissions made by individual subjects. Participants who do not submit solutions simply do not appear. The left panel, showing subjects working under the No Intermediate Disclosures regime quite clearly shows individual subjects making submissions whose quality sometimes increases (upward trajectory between dots) and sometimes decreases (downward trajectory between dots), but generally tends to improve over time. This gyrating trend is consistent with trial-and-error learning–itself consistent with parallel, independent experimentation. Also consistent with the hypothesis is apparent uncorrelatedness across participants' trajectories. Some start high, others low, and with no clear pattern of relatedness among gyrations up and down. These patterns readily demonstrate Prediction 2.

Qualitatively contrasting with these aforementioned patterns, subjects working under Intermediate Disclosures quite clearly are engaging in an utterly different sort of cumulative learning process. In the second panel of Figure 1 we do not see gyrating up-and-down trajectories, but rather a considerably smooth ascent made by individuals and also the maximal frontier of the group as a whole. The trajectories of individuals are also far more clustered around the maximal frontier, particularly as greater time passes (apart from a handful of lower outliers achieving low scores). Under the No Intermediate Disclosures

[8] Reviewing these patterns, it is immediately clear there is also much scope for any number of subtle dynamic games and interactions, such as possible strategic delays, strategic signaling and so on. Discerning these more subtle effects goes beyond the scope of this particular study; however, these are surely important topics worthy of deeper investigation.

policy in the first panel, we see the distribution of scores is instead notably more uniformly and evenly dispersed across the vertical performance spectrum. These patterns readily fall in line with Prediction 3.

<Fig. 1.>

Also to appreciate differences in search patterns, we examine the technical approaches and directions of search pursued under the different disclosure policies. Consistent with the appearance of more convergent, cumulative learning under the Intermediate Disclosures policy in Figure 1, we find there there to be many fewer technical approaches or search directions initiated, and less experimentation under the Intermediate Disclosures policy than under the No Intermediate Disclosures policy. This can be see in Figure 2. As a proxy or indication of unique technical approaches, we count the number of unique combinations of techniques within each submitted solution (among the 10 elemental techniques that were coded, as described in Section 3). Under the policy of No Intermediate Disclosures, we find 27 unique combinations were tried over the two week experiment; whereas, under the Intermediate Disclosures policy, just 19 unique approaches were attempted, 30% fewer. The two trend lines of accumulated experiments across novel approaches never cross. These patterns too are consistent with Hypotheses 2 and 3.

Further corroborating these patterns of lower expansiveness of search under Intermediate Disclosures, the Herfindahl Index measuring the concentration of solutions across different approaches is also 52% higher, at 0.149 in Intermediate Disclosures versus 0.0986 in No Intermediate Disclosures. Also consistent with less experimentation and diversity in Intermediate Disclosures, final solutions appeared in three programming languages in that regime (C#, C++, and Java), whereas in No Intermediate Disclosures 8.7% of final solutions also came from two additional languages (Python and Visual Basic). These patterns are consistent with the wider range of experimentation with No Intermediate Disclosures as per Prediction 2.

<Fig. 2.>

5.3 Individual Subject Innovative Search Behaviour

Apart from the broad patterns at the population level described in the earlier subsection, here we examine evidence at the level of individual subjects' search behaviour and choices. Despite the *collective* number of technical approaches being 30% lower under the Intermediate Disclosures policy, the data suggest a more nuanced set of distinctions at the level of *individual* participants. In at least one sense, subjects working under Intermediate Disclosures were individually relatively active: over all of their submissions, participants in the Open Disclosure regime tried a greater number of elemental techniques, 0.8 more (or 1.3 more, when controlling for skill and numbers of submissions made, significant at $p = 1\%$).

Participants in Intermediate Disclosures also used a greater number of elemental optimization techniques, on average, in their final solutions. However, despite each being aware of a greater number of elemental techniques, they tried our a few number of unique combinations of techniques than did those in No Intermediate Disclosures, .3 fewer (or 0.6 fewer, when controlling for skill, numbers of submissions made, and numbers of individual techniques over which they could experiment, significant at $p = 1\%$). Therefore, not only were there fewer subjects actively participating under Intermediate Disclosures, those who did enter engaged in less experimentation. The high numbers of techniques and low levels of experimentation are consistent with both greater degrees of ex-post reuse of knowledge (as predicted by the incentives-versus-reuse tradeoff) and convergent experimentation (as predicted by differences in the innovation search process).

Consistent with our predictions that intermediate disclosures should go further to alter the direction of search and the innovative process, there is also evidence of greater "targetedness" and convergence of efforts. For example, despite having fewer active participants under Intermediate Disclosures, subjects under Intermediate Disclosures also even came up with fewer *novel* combinations of approaches, .02 fewer (or 0.6 fewer, when controlling for skill, numbers of submissions made, and numbers of individual techniques over which they could experiment, significant at $p = 1\%$).

More striking is a strong tendency to implement only high-potential technical approaches. For example, if we roughly proxy for the "potential" of different technical approaches or combinations of techniques as the top score achieved using a given technique across the entire experiment, we find that the approaches deployed under the Intermediate Disclosures policy each above the median (including the handful of submissions that achieved minimal scores). These patterns illustrate distinct qualitative choices in which technical approaches to implement. These patterns indicate that rather than just a boost in reuse per se, participants are actively interpreting the array of possibilities and coordinating efforts to what appear to be high-potential approaches.

Individual level patterns under the No Intermediate Disclosures policy also exhibit differences in the innovative search process. This begins, of course, with the opposite finding to each of the comments above. For example, participants in this case collectively experimented over a wider set of technical approaches, despite individually trying out and deploying fewer elemental techniques within their solutions. Participants also individually experimented over a greater number of technical approaches or combinations of techniques, despite having fewer techniques to work with. The higher collective number of approaches experimented over is therefore not only a result of having greater entry and greater effort, activity and incentives to innovate per participant–it is that there is also a greater level of experimentation per participant. With a greater level of parallel experimentation, however, came a larger proportion of low-potential approaches having been tried.

5.4 Robustness

The wider experimentation and greater number of technical approaches realized under the No Intermediate Disclosures policy could, in principle, relate to a wide range of explanations. For example, greater experimentation could have simply been driven by higher effort, as predicted by the incentives-versus-reuse account. However, as above, earlier differences in experimentation held whether controlling for our measure of effort (number of submissions) or not.

Greater experimentation might have also been created by non-random sorting into active participation in the development process. Recall, that our research design matched the set of prospective entrants under each disclosure policy, not entrants and active participants. A first indication that sorting and selection cannot account for the earlier results is that differences in technical approaches were reported earlier both unconditionally and conditional on the skill level of participants. To provide still greater assurance that sorting and selection are not driving the earlier results in some more subtle fashion, we investigated the types of individuals who chose to actively enter and participate under the different regimes and found little difference in the compositional distribution of entrants across observable attributes (just differences in absolute numbers who chose to enter and actively participate). We found no statistical or meaningful difference in average skills or distribution of skills (either via testing of first and second moments or application of a Kolmogorov-Smirnov test of differences (Justel et al., 1997). We then also assured that alternative measures of skills–including the year of joining the TopCoder platform and total number of prior events similarly did not statistically differ. We also verified there were no statistically significant differences in technical interests or in country of origin indicated by those choosing to actively enter and participate in either regime.[9]

Finally, as earlier mentioned, our research design strategy of maximising the size of pools of prospective entrants under the different regimes came with the cost of minimising replication. To offer some level of assurance that the results studied above are not somehow

[9]As regards technical interests, TopCoder collected data on the stated main technical interest of participants from 45% of subjects prior to the experiment, at the time at which subjects originally joined the platform. These include: Broadband; Data, Voice, Video Convergence; Game Software Development; Graphic Design; Handheld; Networking; Security; (General) Software Development; Web; and Wireless. As regards country data, TopCoder collected country data on 100% of members. We tested differences in the distribution across top countries–India, USA, Russia, China–along with 65 "Other" countries, given the distribution was very sparse across these other countries.

eccentric and reflect aberrant noise rather than patterns we would expect if we ran the experiment over more trials, we concurrently ran a "Mixed" policy regime. In that regime, there were no intermediate disclosures in the first week, but intermediate disclosures in the second. The results in the mixed regime quite clearly corroborate the earlier patterns. Both the descriptive patterns of submissions and scores, and development of novel combinations of techniques appear to be "between" those patterns documented in Figures 1 and 2. For example, in the first week the Mixed regime produces a series of gyrating trajectories, similar to what is observed in the first panel of Figure 1. In the second week, the Mixed regime produces a much more laminar set of trajectories, akin to those observed in the second panel of Figure 2. Further corroborating the results, the trajectory of the number of novel approaches tried sits between the trajectories for the other two regimes, described in Figure 2. While the results in the Mixed regime hardly confirm that this regime should be regarded as simply some sort of average of the two main comparison policies in our main analysis, the Mixed regime does provide some assurance that the earlier results are not somehow eccentric.

6 Summary and Conceptual and Policy Implications

In this paper, we reported on results from a field experiment designed to investigate the effects of intermediate versus final disclosure policies applied to similar populations of subjects on innovation outcomes. In addition to hypothesizing a classical incentives-versus-reuse tradeoff, our main interest was to investigate whether intermediate and final disclosures led to qualitatively different innovative search processes, with final disclosures leading to parallel experimentation in the population of innovators and intermediate disclosures initiating a process of cumulative innovation focused on known solution pathways. We were especially motivated to investigate these differences in disclosure policies given these sorts of disclosure policy differences are most salient across entirely different innovation systems, as in Figure 3, suggesting the possibility these differences might somehow be inherent to their comparative advantages in supporting under different conditions or contingencies.

<Figure 3>

Subjects in our experiment created solutions to a challenging bioinformatics problem that both industrial and academic labs currently face, and one that has been subject to a process of cumulative innovation outside of our experiment with significant investments by public and private sector actors. Our subjects possessed relevant aptitudes and a mix of skills to address the problem at hand–and created working solutions that exceeded the performance of the existing state of the art (see Lakhani et al. 2013). Under controlled conditions, independent groups of subjects worked under alternative disclosure policies. Our analysis was not designed to reveal subtle, or difficult-to-detect "effects" that might explain say some fraction of variation in innovative output. Rather, our goal was to observe whether readily apparent differences would emerge upon implementing relevant comparison groups and devising appropriate measures to discern the innovative search process.

Our analysis began by first confirming the presence of the classically theorised tensions between ex-ante incentive provision and ex-post productive reuse. In relation to this secondary finding, it is worth noting that existing empirical research tends to just focus on just ex post reuse conditional on the appearance of an innovation (e.g., Murray et al., 2009; Boudreau 2010; Furman and Stern, 2011; Galasso and Schankerman; Williams, 2013). This necessarily provides an optimistic suggestion of effects of disclosures, as greater disclosures and freer access should indeed tend to increase knowledge reuse, as this is just one side of the tradeoff. Here, in our results, we examine both sides of the tradeoff at once, showing starkly lower effort and participation with greater disclosure, and also showing that the stock of knowledge (that is more freely reused) is itself substantially smaller with greater disclosures.

The main objective of our investigation was to study patterns of innovative search arising from intermediate and final disclosure policies. We readily observe striking results even in plotting aggregate, population-level patterns. Intermediate disclosures led to a highly "laminar" or smooth process of ascent in performance scores. Whereas the final disclosure policy individual performance trajectories were erratic and mostly uncorrelated with those of others across the population. Performance under the intermediate disclosure policy converged to the maximal frontier with multiple innovators achieving similarly high performance. Under the final disclosure policy, solutions were much more evenly dispersed across the performance spectrum, from high quality to low quality. Further, final disclosures led to significantly greater number of distinct technical approaches being attempted attempted.

Therefore, the findings here add an additional predictive insight beyond the usual incentives-versus-reuse tradeoff characterisation. Intermediate disclosures not only led to an increase in reuse. Because participants actively interpreted the array of possibilities with the support of the information and signals generated by others' activities, development activity became coordinated on highest-potential existing technical approaches. This patterning of the direction of experimentation might, in principle, either systematically increase or decrease performance, depending on the existing stock of knowledge on which reuse occurs. Greater convergence of search towards higher performing solutions will indeed incline attention towards superior existing approaches, conditional on the stock of knowledge. However, if the existing stock of knowledge does not somehow lead innovative efforts towards globally optimal approaches, an intermediate disclosure policy should be more likely to lead a population of innovators to get locked-in to suboptimal technical approach or trajectory, all else being equal.

Opposite sorts of arguments apply to the case of final disclosures and restricting intermediate disclosures. On the basis of the incentives-versus-reuse characterisation, on its own, we should expect that restricting intermediate disclosures should generate higher experimentation, if only on the basis of higher incentives alone. However, just as those working under intermediate disclosures experience disproportionately higher average productivity (conditional on the stock of knowledge and levels of effort) than what would otherwise be predicted, those working under no intermediate disclosures are likely to achieve lower average outcomes than what would otherwise be predicted. Although efforts are high, experimentation proceeds in a range of directions, including many low-potential directions. The wider expanse of experimentation, however, diminishes the probability that the system overall gets stuck on locally optimal solutions.

The above findings related to the patterning of innovation and search processes, our findings contribute to canonical characterisations of innovation processes, which have tended to emphasise the importance of incentives and knowledge reuse, particularly in formal modelling (Romer, 1991; Green and Scotchmer, 1995) while corroborating theories on the tendency to convergent pathways and paradigms (e.g., Kuhn 1962, Abernathy and Utterback 1978, Sahal 1981, Dosi 1982). Our work also relates to separate traditions studying the role and importance of independent parallel search paths in innovation have tended to emphasise the effect of restricting market signals and information (esp. Nelson, 1962; Scherer 2011), and the effects of incentives (Terwiesch and Xu, 2008; Boudreau et al., 2011). Here we show each of these effects within a single empirical context and, in so doing, clarifying resulting *collective* or *population-level* patterns of innovative search initiated by alternative disclosure policies. These findings might also be interpreted as inviting greater bridges to be built between innovation policy literatures to those beginning to develop formal models of search processes, modelling "rugged landscapes" and solution spaces (e.g., Levinthal,1997**;**Csaszar and Siggelkow, 2010; Rivkin and Siggelkow, 2003, 2005)).

The Comparative Advantages of Different Open Innovation Systems In documenting fundamental differences in the collective search process initiated by alternative disclosure policies, our results also provide some clues as to the relative strengths and weaknesses of different innovation systems–and the sorts of problems and contexts we should see them supporting.

Intermediate disclosures make better, more targeted, coordinated, efficient use of existing knowledge, but–all else being equal–they suffer the important pitfall of being especially susceptible to path dependence and converging on a suboptimal solution paths. It should not therefore be surprising that in many cases, we see intermediate disclosure policies used in contexts in which there already exists a large stock of knowledge and a mature history of development, helping minimize these pitfalls. The problem of path dependence might also be mitigated to the extent that intermediate disclosure policies are implemented where the sheer vastness and diversity of the pool of contributors injects a variety of perspectives and thus reduces the hazard of locking into suboptimal solutions (Arthur, 1989; Page, 2006). For example, open source software projects benefit from decades of development experience since software developed into a separate and distinct area of technical development in the 1950s; they also draw from the knowledge and experience from a global population of open source developers much larger than the internal software teams of the world's largest companies.

A second pitfall of policies of intermediate disclosures is the potential harm they do to the incentives to invest in innovation. This problem may be partially offset by the types of problems on which intermediate disclosure systems specialise. For example, open data systems are often meant to foster different sorts of solutions on the basis of a common knowledge asset, allowing differentiation in the application to which upstream knowledge resources are applied. Open source software arguably benefits at this stage to a greater degree from "many eyeballs" rather than the heroic contributions of any one genius developer. Further, successful intermediate disclosure systems, as open source software, implement compensating mechanisms such as attribution and reputation systems and informal norms of sharing and collective prosocial orientation to continue to encourage highly talented developers to continue to make contributions. In this case, developers (or their employing organisations) may experience high incentives to make contributions simply because they benefit from own use (von Hippel 2005) or seek other benefits (Lakhani and von Hippel, 2003; Lakhani and Wolf, 2005; Roberts et al., 2006). Inasmuch as intermediate disclosures reduce innovators protections and curb incentives, there is also a clear requirement that the necessary scale of investment remain small. Indeed the general work on modularity in technological systems (Baldwin and Clark, 2000) and its specific application to open source software (Baldwin and

Clark, 2006; MacCormack et al., 2006, 2012) illustrates that such systems may indeed be designed to be more modular and hence to lower the fixed cost of individual participation and contribution.

Restricting intermediate disclosures and admitting just final disclosures has the advantage of promoting experimentation–but with the weakness of high cost and split attention and direction of research activity. It is for this reason that we may often see the use of systems especially geared to promote large scale independent parallel experimentation–as in crowdsourcing contests and grand innovation challenges–invoked to address challenging innovation problems–where there is a premium on experimentation that warrants the cost (Jeppesen and Lakhani, 2010; Boudreau et al. 2011; Murray et al. 2013). Contests have also now been designed with compensating mechanisms to minimise their pitfall of incurring high costs, by often stimulating participation through the implicit incentives provided by contest participation rather than always explicit cash prize incentives (Boudreau et al. 2011; Murray et al. 2013). Open "apps" platforms have similarly been designed to maintain many protections and secrecies and parallel experimentation to generate extraordinary variety of product offerings (Boudreau, 2012). In addition, our entrepreneurial innovation system, whereby start-up firms seeking private funding and then attempt to develop proprietary technologies and protected business models might also be interpreted in this light. Inasmuch as such businesses are attempting to truly innovate something new to the world, the innovation challenge is non-trivial and society benefits from having multiple attempts reflecting idiosyncratic beliefs and capabilities rather than convergent approaches.

Implications for Firm Strategy While each quadrant in Figure 3 represents a distinctive perspective and a set of policy prescriptions for participants in the particular system, it is also interesting to consider choices made by various actors to simultaneously participate in multiple quadrants. For example, IBM as an emblematic innovative firm makes significant annual investments to increase its patent portfolio, while simultaneously encouraging basic research and development with industrial and academic partners and is one of the most active participants in the open source software projects (Capek et al., 2005) . Underlying their decision to participate in three distinctive innovation system instead of simply focussing on internal innovation efforts is the recognition that the nature of problems, the distribution of knowledge, the availability of participants and partners and the ability to extract profits (either directly through sales of goods or through complementary assets) can strategically shape their choices of the disclosure and incentive regimes they choose. Business case studies have shown that large firms like Apple and LEGO are now making such choices routinely (Lakhani et al., 2013) and many entrepreneurial firms, in areas as distinctive as graphic design (e.g., Threadless), automotives (e.g., Local Motors) and consumer product development (e.g., Quirky), are "born open," with business models that straddle all three quadrants of Figure 3.

Innovation policy makers and funders of national innovation initiatives are also faced with similar options with regards to the allocation of resources towards innovative activity and their policies for the disclosure of the outputs. While various open access policies ensure full availability of funded research as it is completed, the embrace of the Bermuda Principles by the NIH and other national funding authorities recognised that the nature of the task in front of them (sequencing the human genome) and the presence of private competition marked an important departure towards a regime of intermediate disclosures that coordinated the efforts of thousands of scientists. Nevertheless, while the Human Genome Project was a significant and important scientific success, there has not been an institutionalisation of such practices within academic sciences, thus hinting at the underlying tension between our current system of incentives and the reuse of knowledge.

Final Conclusion We hope that our paper provides a useful lens unto the ongoing debates between "closed" and "open" innovation systems and the design of institutions supporting innovation more generally. Rather than continuing the debate on the various merits of each system, our intended approach here was to study one essential difference across different innovations systems, intermediate versus final disclosure policies and to examine the effect of such policies in a field experimental setting. While there has been significant enthusiasm for innovation systems that embrace intermediate disclosure within the academic research and in the popular press, we provide experimental evidence, with appropriate counterfactuals, to show that there are both costs and benefits need to be considered. Each approach has its comparative advantages. Systems implementing intermediate disclosure make efficient and deliberate use of existing knowledge and implement a degree of coordination of efforts. However, all else being equal, innovators working under such systems will experience lower economic incentives to make investments and it will also curtail the extent and expanse of innovative search and technical experimentation.

REFERENCES

Abernathy, W. J., Rosenbloom, R.S., 1969. Parallel strategies in development projects. Management Science 15:10, 486–505.

Abernathy, W.J., Utterback, J.M., 1978. Patterns of industrial innovation. Technology Review 80, 40–47.

Allen, R., 1983. Collective Invention. Journal of Economic Behavior and Organization 4, 1– 24.

Altschul, S., Gish, W., Miller, W., Myers, E., Lipman. D., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Arrow, K.,1962. Economic Welfare and the Allocation of Resources for Innovation, in: The Rate and Direction of Inventive Activity, Nelson R., (ed), Princeton University Press, Princeton, NJ.

Arora, A., Fosfori, A., Gambardella, A., 2004. Markets for Technology: The Economics of Innovation and Corporate Strategy. MIT Press, Cambridge, Mass.

Arthur, B.W., 1994. Increasing Returns and Path Dependence in the Economy. University of Michigan Press, Ann Arbor.

Baldwin, C., Clark, K., 2000. Design Rules. The Power of Modularity. The MIT Press, Cambridge, MA.

Baldwin, C., Clark, K., 2006. The Architecture of Participation: Does Code Architecture Mitigate Free Riding in the Open Source Development Model?. Management Science 52:7, 1116-1127.

Baldwin, C., Hienerth, C., and von Hippel, E., 2006. How User Innovations Become Commercial Products: A Theoretical Investigation and a Case Study. Research Policy 35, no. 9.

Bessen, J., Maskin, E., 2009. Sequential innovation, patents, and imitation. RAND Journal of Economics 40, 611-635.

Bijker, W. E., Hughes, T.P., Pinch, T.J., 1987. The Social Construction of Technological

Systems: New Directions in the Sociology and History of Technology. MIT Press, Cambridge, Mass.

Boudreau, K., 2010. Open platform strategies and innovation: Granting access vs. devolving control. Management Science 56, 1849–1872.

Boudreau, K., 2012. Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software 'Apps' Developers and Patterns of Innovation. Organization Science 23, 1409-1427.

Boudreau, K., Lacetera, N., Lakhani, K.R., 2011. Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. Management Science 57, 843-863.

Caballero, R. J., Jaffe, A. B., 1993. How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth, NBER Chapters, in: NBER Macroeconomics Annual 1993, Volume 8, pages 15-86 National Bureau of Economic Research, Inc.

Capek, P.G., Frank, S.P., Gerdt, S., Shields, D., 2005. A history of IBM's open-source involvement and strategy. IBM Systems Journal 44:2, 249-258.

Chesbrough, H., 2003. Open Innovation. Harvard Business Press, Boston, Mass.

Cockburn, I., Stern, S., 2010. Finding the Endless Frontier: Lessons from the Life Sciences Innovation System for Technology Policy. Capitalism and Society 5:1, article 1.

Cohen, W. M., Nelson, R. R., Walsh, J. P., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why US Firms Patent (or Not). NBER Working Paper No. 7552.

Contreras, J., 2011. Bermuda's Legacy: Policy, Patents, and the Design of the Genome Commons. Minnesota Journal of Law, Science & Technology 12, 61-125.

Csaszar, F., Siggelkow, N., 2010. How much to copy? Determinants of Effective Imitation Breadth. Organization Science 21, 661-676.

Dahan, E., Mendelson, H., 2001. An extreme-value model of concept testing. Management Science 47, 102–116.

Dasgupta, P., David, P., 1994. Toward a new economics of science. Research Policy, 23, 487-521.

David, P., 1985. Clio and the Economics of QWERTY. American Economic Review 75, 332–337.

Deogirikar, A., Stebbin, M., 2013. Seeking Outstanding 'Open Science' Champions of Change. White House Office of Science and Technology Policy Press Release, URL: http://www.whitehouse.gov/blog/2013/05/07/seeking-outstanding-open-science-champions-change.

Dosi, G., 1982. Technological paradigms and technological trajectories. Research Policy 11, 147-162.

Eisenmann, T. R., Parker, G., Van Alstyne, M., 2009. Opening Platforms: When, How and Why? Chap. 6 in Platforms, Markets and Innovation. Paperback (ed) Edited by Annabelle Gawer. Cheltenham, U.K. and Edward Elgar Publishing, Northampton, Mass.

Erat, S., Krishnan, V., 2011. Managing Delegated Search Over Design Spaces. Management Science 58:3, 606-623.

Farrell, J., 1987. Cheap talk, coordination, and entry. The RAND Journal of Economics 18, 34-39.

Fauchart, E., von Hippel, E., 2008. Norms-Based Intellectual Property Systems: The Case of French Chefs. Organization Science 19, 187-201.

Fleming, L., Sorenson, O., 2001. Technology as a complex adaptive system: evidence from patent data. Research Policy 30, 1019–1039.

Fleming, L., Sorenson, O., 2004. Science as a map in technological search. Strategic Management Journal 25, 909–928.

Furman, J., Stern, S., 2011. Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research. American Economic Review 101, 1933–1963.

Hounshell, D., 1985. From the American System to Mass Production, 1800-1932: The development of manufacturing technology in the United States. Johns Hopkins University

Press. Baltimore, MD.

Jeppesen, L., Laursen, K., 2009. The Role of Lead Users in Knowledge Sharing. Research Policy 38, 1582-1589.

Jeppesen, L., Lakhani, K.R., 2010. Marginality and problem- solving effectiveness in broadcast search. Organization Science, 21:5, 1016–1033.

Galasso, A., Schankerman, M., 2013. Patents and Cumulative Innovation: Causal Evidence from the Courts. CEP Discussion Paper No. CEPDP1205.

Gans, J., Murray, F., 2012. Funding Scientific Knowledge: Selection, Disclosure and the Public-Private Portfolio Ch. 1 in: The Rate and Direction of Inventive Activity Revisited, J. Lerner and S. Stern, (eds) University of Chicago Press, Chicago.

Green, J., Scotchmer, S., 1995. On the Division of Profit in Sequential Innovation. The RAND Journal of Economics 26, 20-33.

Gowers, T., Michael Nielsen, M., 2009. Massively collaborative mathematics. Nature 461, 879–881.

Haeussler, C., Jiang, L., Thursby, J., Thursby, M., 2009. Specific and General Information Sharing Among Academic Scientists. NBER Working Paper Series 15315.

Harhoff, D., Henkel,K., E Von Hippel, E., 2003. Profiting from voluntary information spillovers: how users benefit by freely revealing their innovations. Research policy 32:10, 1753-1769.

Heller, M., Eisenberg, R., 1998. Can patents deter innovation? The anticommons in biomedical research. Science 280, 698-701.

Hotelling, H., 1929. Stability in Competition. The Economic Journal 39, 41-57.

Jaffe, A. B., Trajtenberg, M., Fogarty, M., 2000. Knowledge spillovers and patent citations: Evidence from a survey of inventors. American Economic Review Papers and Proceedings 90, 2: 215-218.

Jovanovic, B., 1984. Matching, Turnover, & Unemployment. Journal of Political Economy 92, 108-122.

Kitch, E., 1977. The Nature and Function of the Patent System. Journal of Law and Economics 20, 265-290.

Klein, B., Lerner, A. V., Murphy, K. M., 2002. The Economics of Copyright "Fair Use" in a Networked World. American Economic Review, 92:2, 205-208.

Klemperer, P., 1990. How Broad Should the Scope of Patent Protection Be? RAND Journal of Economics, 21:1, 113-131.

Kuhn, T., 1962. The Structure of Scientific Revolutions. University of Chicago Press, Chicago.

Laursen, K., Salter, A., 2006. Open for Innovation: The Role of Openness in Explaining Innovation Performance Among U.K. Manufacturing Firms. Strategic Management Journal 27, 131-150.

Lerner, J., Schankerman, M., 2010. The Comingled Code: Open Source and Economic Development. MIT Press: Cambridge.

Lessig, L., 1999. Code and Other Laws of Cyberspace. Basic Books: New York.

Levinthal, D.A., 2007. Adaptation on Rugged Landscapes. Management Science 43, 934-950.

Lakhani, K.R., Boudreau, K., Loh P-R., Backstrom, L., Baldwin, C., Lonstein, E., Lydon, M., MacCormack, A., Arnaout, R., Guinan, E., 2013. Prize-based experiments can provide solutions to computational biology problems. Nature Biotechnology 31, 108-111.

Lakhani, K.R. Lifshitz-Assaf, H., Tushman, M., 2013. Open Innovation and Organizational Boundaries: Task Decomposition, Knowledge Distribution and the Locus of Innovation in: Handbook of Economic Organization: Integrating Economic and Organization Theory, (ed), A. Grandori: Edward Elgar Publishing, Northampton, Mass., pp. 355-382.

Lakhani, K.R., Wolf, R., 2005. Why hackers do what they do: Understanding motivation and effort in free/open source software projects in: H. Feller, B. Fitzgerald, S. A. Hissam, & K. R. Lakhani (Eds.), Perspectives on free and open source software. MIT Press, Cambridge, Mass.

Lerner, J., 2006. 150 Years of Patent Office Practice. Law and Institutions. American Law and Economics Review 7, 112–143.

Lerner, J., Tirole, J., 2005. The economics of technology sharing: Open source and beyond. The Journal of Economic Perspectives 19, 99-120.

Lessig, L., 2009. Remix: Making Art and Commerce Thrive in the Hybrid Economy. Penguin Books, New York.

Levinthal, D.A., 1997. Adaptation on Rugged Landscapes. Management Science 43, 934 - 950.

Lin, T., 2012. Cracking Open the Scientific Process. The New York Times.

Maas, H., Wagenmakers. E-J., 2005. A Psychometric Analysis of Chess Expertise. The American Journal of Psychology 118, 29-60.

MacCormack, A., Rusnak, J. Baldwin, C., 2006. Exploring the Structure of Complex Software Designs: An Empirical Study of Open Source and Proprietary Code. Management Science 52:7, 1015-1030.

Mansfield, E., 1985. How rapidly does new industrial technology leak out? Journal of Industrial Economics 34, 217–223.

Marburger, J., 2005. Wanted: Better Benchmarks. Science 308,1087.

March, J.G., 1991. Exploration and exploitation in organizational learning', Organization Science, 2:1, 71–87.

Mazzoleni, R., Nelson, R., 1998. The benefits and costs of strong patent protection: a contribution to the current debate. Research Policy, Elsevier 27:3, 273-284.

Meyer, P. B., 2013. The Airplane as an Open Source Invention. Revue économique, 64:1, 115-132.

Mokyr, J., 2002. The Gifts of Athena: Historical Origins of the Knowledge Economy. Princeton University Press, Princeton, NJ.

Moon, S., 2011. How does the management of research impact the disclosure of knowledge? Evidence from scientific publications and patenting behavior. Economics of Innovation and New Technology 20, 1–32.

Moser, P., 2005. How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World Fairs. The American Economic Review 95,1214-1236.

Mukherjee, A., Stern, S., 2009. Disclosure or secrecy? The dynamics of Open Science. International Journal of Industrial Innovation 27, 459-462,

Murman, J., Tushman, P., Tushman, M., 1998. Dominant Designs, Technology Cycles, and Organizational Outcome. Research in Organizational Behavior 20.

Murray, F. O'Mahony, S., 2007. Exploring the foundations of cumulative innovation: implications for organization science. Organization Science 18:6, 1006–1021.

Murray, F., Stern, S., 2007. Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis. Journal of Economic Behavior and Organization 63, 648–687.

Murray, F., Aghion, P., Dewatripont, M. M., Kolev, J., Stern, S., 2009. Of Mice and Academics: Examining the Effect of Openness on Innovation. NBER Working Paper Series 14819.

Murray, F., Stern, S., Campbell, G., MacCormack, A., 2012. Grand Innovation Prizes: A Theoretical, Normative and Empirical Evaluation. Research Policy, 41(10):1779-1792.

Nordhaus, W. D., 1972. The Optimal Life of a Patent. The American Economic Review, 62:3, 428-431.

Nelkin, D., 1982. Intellectual property: the control of scientific information. Science 216, 704-708.

Nelson, R., 1959. The simple economics of basic research. Journal of Political Economy 67, 297-306.

Nelson, R., 1961. Uncertainty, Learning, and the Economics of Parallel Research and Development. Review of Economics and Statistics 43, 351-368.

Newell, Allen, and H.A. Simon. 1972. Human Problem Solving. Engelwood Cliffs, New Jersey: Prentice-Hall INC.

Nickerson, J. A., Zenger, T. A., 2004. A knowledge-based theory of the firm - the problem solving perspective. Organization Science 15:617-632.

Nielsen, M., 2011. Reinventing Discovery: The New Era of Networked Scienc. Princeton University Press, Princeton, NJ.

O'Mahony, S., 2003. Guarding the commons: how community managed software projects protect their work. Research Policy 32, 1179-1198.

Osterloh, M., Rota, S., 2007. Open source software development: just another case of collective invention? Research Policy 26:2, 157-171.

Page, S., 2007. The Difference. Princeton University Press, Princeton, NJ.

Raasch, C., von Hippel, E., 2013. Innovation Process Benefits: The Journey as Reward. Sloan Management Review 33-39.

Raymond, E., 1999. The Cathedral and the Bazaar. Sebastapol: O'Reily and Associates.

Rivkin, J. W., Siggelkow, N., 2003. Balancing Search and Stability: Interdependencies Among Elements of Organizational Design. Management Science 49, 290–311.

Rivkin, J., Siggelkow, S., 2007. Patterned Interactions in Complex Systems: Implications for Exploration. Management Science 53, 1068-1085.

Roberts, J., Hann, I., Slaughter, S., 2006. Motivations, Participation and Performance in Open Source Software Development. Management Science 52, 984-999.

Romer, P., 1990. Endogenous Technological Change. Journal of Political Economy 98:5. Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems. University of Chicago Press, Chicago. pp. S71-S102.

Royal Society., 2012. Science as an open enterprise. The Royal Society Science Policy Centre report 02/12.

Rosenberg, N., 1976. Perspectives on Technology. Cambridge University Press, Cambridge.

Rosenberg, N., 1984. Inside the Black Box: Technology and Economics. Cambridge University Press.

Rosenkopf, L., Nerkar, A., 2001. "Beyond local search: Boundary-spanning, exploration and impact in the optical disk industry." Strategic Management Journal 22:287-306.

Rysman, M., Simcoe, T., 2008. Patents and the Performance of Voluntary Standard-Setting Organizations. Management Science 54, 1920-1934.

Sahal, D., 1981. Patterns of technological innovation. Addison-Wesley Publishing Company, Reading, Mass.

Salop, S., 1979. Monopolistic Competition with Outside Goods. The Bell Journal of Economics 10, 141-156.

Scherer F.M., 2011. Parallel R&D Paths Revisited. Harvard Kennedy School Faculty Research Working Paper Series No. RWP11-022.

Scherer F.M., Harhoff, D., Kukies, J., 2000. Uncertainty and the Size Distribution of Rewards from Technological Innovation. Journal of Evolutionary Economics 10, 175-200.

Scotchmer, S., 1991. Standing on the Shoulders of Giants: Cumulative Research and the Patent Law, Journal of Economic Perspectives 5(1), 29-41.

Scotchmer, S., 2004. Innovation and Incentives. The MIT Press, Cambridge, Mass.

Simcoe, T., 2007. Explaining the Increase in Intellectual Property Disclosure in: The Standards Edge: The Golden Mean. Bolin Group

Simon, H. A. 1962. The architecture of complexity. Proc. Amer. Philos. Soc. 106(6) 467–482.

Sorenson, O., Rivkin, J.W., Fleming, L., 2006. Complexity, networks and knowledge flow. Research Policy, 35, 994-107.

Stephan, P., 1996. The economics of science. Journal of Economic Literature 34, 1199-1235.

Stern, S., 2004. Biological Resource Centers: Knowledge Hubs for the Life Sciences. The Brookings Institution Press, Washington D.C.

Terwiesch, C., Ulrich, K., 2009. Innovation Tournaments: Creating and Selecting Exceptional Opportunities. Harvard Business School Press, Boston, Mass.

Terwiesch, C., Xu, Y., 2008. Innovation contests, open innovation, and multiagent problem solving. Management Science 54, 1529–1543.

Utterback, J., 1994. Mastering The Dynamics of Innovation. Harvard University Business School Press, Boston, Mass.

Utterback, J., Abernathy, W., 1975. A Dynamic Model of Product and Process Innovation. Omega 3, 639-656.

Utterback, J., Suárez, F., 1993. Innovation, Competition and Industry Structure. Research Policy 22, 1-21.

Varian, H.R., 2005. Copying and Copyright. Journal of Economic Perspectives 19:2, 121-138.

Vincenti, W. G., 1994. The Retractable Airplane Landing Gear and the Northrop "Anomaly": Variation-Selection and the Shaping of Technology. Technology and Culture 35, 1-33.

von Hippel, E., 1987. Cooperation Between Rivals: Informal Know-How Trading. Research Policy 16, 291-302.

von Hippel, E., 1988. The Sources of Innovation. MIT Press, Cambridge, Mass.

von Hippel, E., Krogh, G., 2003. Open Source Software and the 'Private-collective' Innovation Model: Issues for Organization Science. Organization Science 14:2, 209–223.

von Hippel, E., 2005. Democratizing Innovation. MIT Press, Cambridge, Mass.

von Hippel, E., Ogawa, S., de Jong, J.P.J., 2011. The Age of the Consumer-Innovator. MIT Sloan Management Review, Fall.

von Krogh, G., Haefliger, S., Spaeth, S., Wallin, M., 2012. Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development. MIS Quarterly 36, 649-676.

von Tunzelmann, N., Malerba, F., Nightingale, P., Metcalfe, S., 2008. Technological paradigms: past, present and future. Industrial and Corporate Change, 17:3, 467-484. doi: 10.1093/icc/dtn012

West, J., 2003. How open is open enough? Melding proprietary and open source platform strategies. Research Policy 32:7, 1259-1285.

Williams, Heidi., 2013. Intellectual Property Rights and Innovation: Evidence from the Human Genome. Journal of Political Economy 121, 1-27.

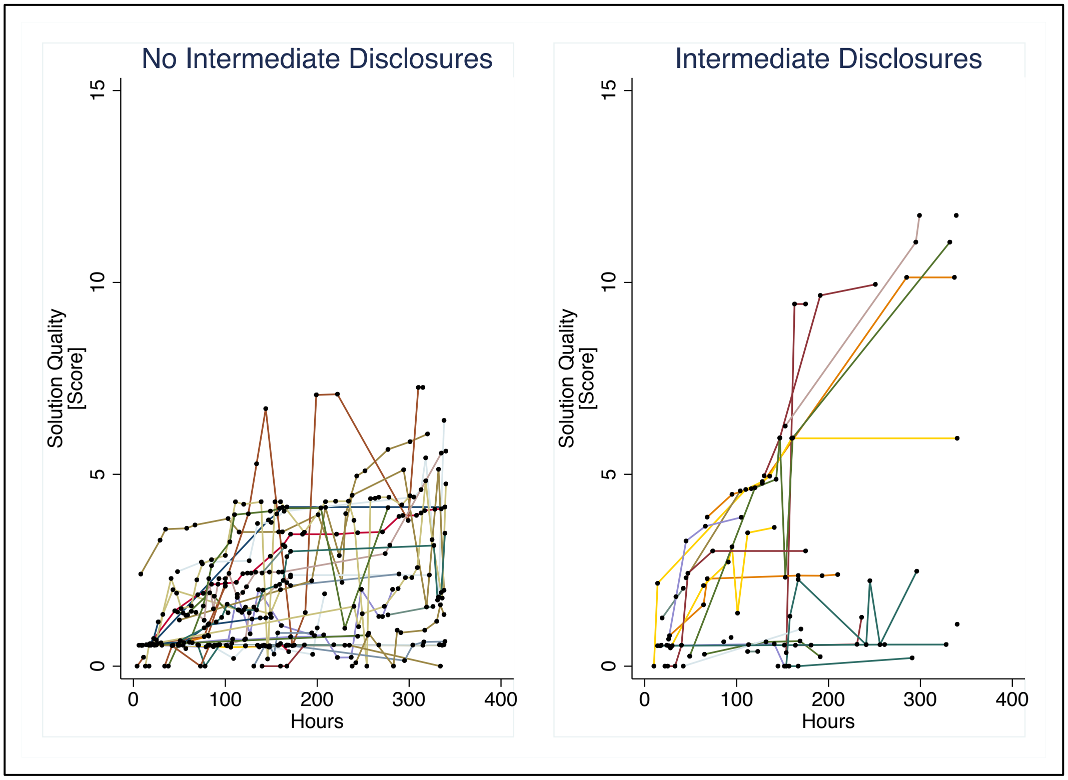

Fig. 1. Performance Trajectories of Individual Participants Under Alternative Disclosure Policies

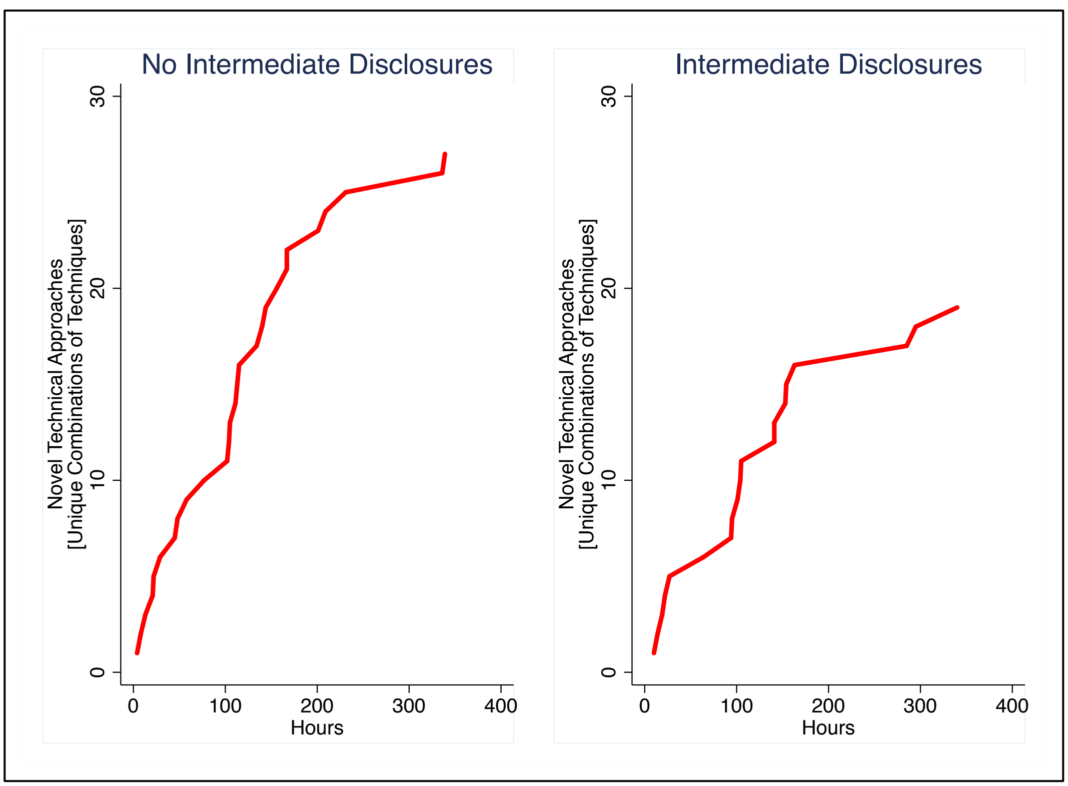

Fig. 2. Total Experimentation and Novel Technical Approaches Tried Under Different Disclosure Policies

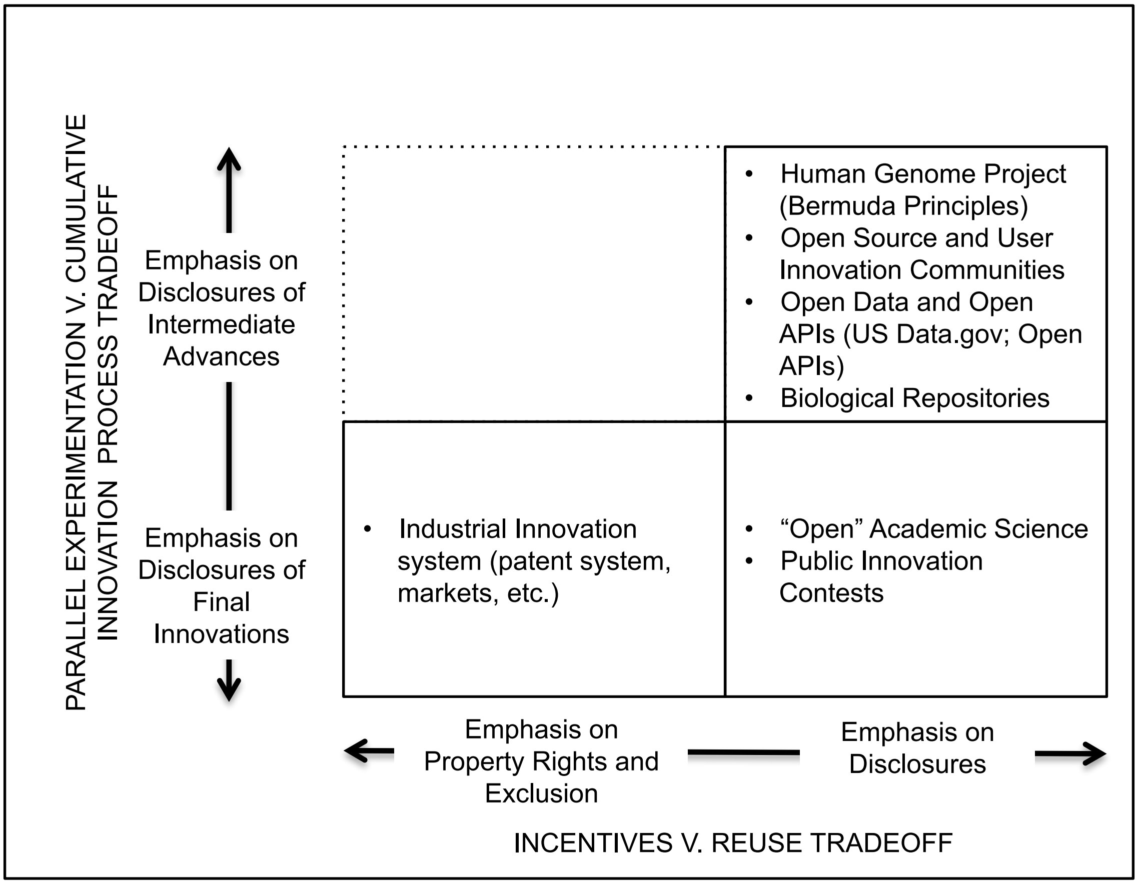

Fig. 3. Broad Differences in Disclosure Policies Across Different Innovation Systems