Teams in R&D: Evidence from US Inventor Data

Jinyoung Kim and Gerald Marschke[1]

December 2016

Abstract:

This paper exploits U.S. patent data and a panel of inventors listed on U.S. patents since 1975 to investigate the determinants of teamwork in industrial R&D. Inventor team size as well as the length of collaboration among team members have increased over the past several decades. The focus of the paper is a test of a model of dynamic team formation where a firm must choose and then over time rebalance a research team's constitution taking into account the gains to specialization, costs of coordination, technological change, and the risks that employee members of the research team will appropriate the firm's intellectual property. We use variation in policy towards noncompete agreements in employment contracts to identify the effect of researcher mobility and IP appropriation on team formation. We find that where researcher job mobility is low, teams tend to be larger and are more likely to repeat. Our evidence suggests that in assembling R&D teams, firms are sensitive to the costs of appropriation and/or coordination.

JEL: J61, D23, J44

Keywords: Teamwork, R&D, Appropriation, Mobility, Team size, Team persistence

[1] Jinyoung Kim, Department of Economics, Korea University; email: jinykim@korea.ac.kr. Gerald Marschke, SUNY Albany, and NBER; email: gerald.marschke@gmail.com. We thank seminar participants at the NBER Productivity Lunch, Binghamton University, the IZA, and Bocconi University as well as participants at the 2010 NBER Summer Institute, Pierre Azoulay, Richard Freeman and Andrew Wang for helpful comments. We also thank Norm Bishara for use of and assistance with his state-level index of noncompete enforcement and Sifan Zhou for assistance with extending the years of coverage of this index. Kim was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A3A2923769). Marschke is grateful to the Science & Engineering Workforce Project (SEWP), and to the Sloan Foundation for research support, as well as for the hospitality of the Labor and Worklife Program at Harvard Law School. Any views or errors are the sole responsibility of the authors.

1. Introduction

Technological innovation increasingly occurs in teams. In 1975, for example, only 42 percent of patents listed multiple inventors and the average number of inventors per patent was 1.6. Today, over two-thirds of US patents granted list multiple inventors and the average number of inventors is 2.6.[2] While the rising importance of teamwork is nearly universal, striking variation exists across technological classes and regions. "Drugs and Chemical" patents averaged nearly one more inventor than "Mechanical" patents in 2003 (3.1 vs. 2.4).[3] Within a technology, inventor team size varies by geography. Minnesota and New Jersey both with significant pharmaceutical and biotechnology industries, yet in New Jersey the average number of inventors on a drug or chemical patent is 25 percent greater (3.54 vs. 2.82) compared to such patents from Minnesota. The inventor team size varies by country as well (see below).

Firms that are unable to field large, diverse teams are arguably at a productivity disadvantage in R&D (see Wuchty, Jones, and Uzzi, 2007), as are firms that are unable to keep their teams intact over a sustained research campaign (Akgun and Lynn, 2002). Continuity in teamwork improves team productivity if for example it takes time for members of a team to learn how to effectively work with one another, or if departing team members leave with key knowledge that is not easily transmitted to remaining team members or their replacements. We show in the patent data that teams vary in longevity and that longer-lived teams are associated with higher-quality innovations.

[2] Scientists working in solitude are also less common. In 1955 only half of academic articles published in science and engineering fields listed two or more authors. In the five years leading up to 2000, however, 80 percent of articles published in science and engineering journals listed multiple authors (Wuchty, Jones, and Uzzi, 2007).

[3] In science team size also varies by field. For example, authorship teams in medicine averaged over four scientists while mathematics averaged less than two for the 1996 to 2000 period (Wuchty, Jones, and Uzzi, 2007).

We also show that teams in R&D persist over longer periods of time and over more projects than in decades past, but, as in team size, team persistence varies across fields and regions.

This paper explores the determinants of team formation in industrial R&D, seeking to explain variation across time, location, and field in both team size and duration. Our point of departure is an economic model of teams (e.g., Becker and Murphy, 1992), in which the optimal size and composition of teams balance the gains to specialization against coordination or information costs. Teams are large when specialization by workers or the gains to collaboration in the technological domain is great, or when coordination costs are low. In this framework, rising team size may be due to falling coordination costs, for example, because of improved strategies that limit free-riding and agency problems, or improvements in communication technologies. Team size may also be increasing because of the rising stock of knowledge; the optimal response by researchers to the rising burden of knowledge may be greater specialization which requires more collaboration (Jones, 2006).

Because we are also interested in understanding team persistence and threats imposed by workers defecting to competitors in future periods, we consider a dynamic variant of the standard, static model. In our model, the firm manages its R&D workforce over multiple periods and in each period the firm and its researchers are hit with technological and human capital productivity shocks that change the optimal mix of skills and therefore workers, and which potentially cause the researchers to depart to compete against the firm. Thus R&D-performing firms face the prospect that their researchers will leave to work for competitors or start up firms of their own that directly compete against them. This risk is suggested by stories appearing in the business press of high tech firms actively encouraging defections among a competitor's technological workforce to assess its technologies (see Kim and Marschke, 2005, and the references therein). Bhide (2000) reports that 71 percent of the firms listed on the Inc 500[4] were built upon replicated or modified ideas developed in their founders' previous employment.[5] To the extent firms face this appropriation cost they have an incentive to reduce the number of research personnel involved in a project. But they also have an incentive to "compartmentalize" their research projects; that is, to spread the research tasks around greater numbers of researchers so that single researchers lack sufficient information to recreate the project on their own.[6] Team formation is thus part of the firm's strategy to protect its IP from its workers. In addition to testing for gains to specialization and coordination costs as determinants of team size, team persistence and team composition, our empirical analysis includes a test for the threat of worker appropriation of IP.

[4] The Inc. 500 is a list of young, fast growing firms published annually by the editors of *Inc*.

[5] See also footnote 1 in Rajan and Zingales (2001).

[6] What we are calling compartmentalization is a practice that exists at least since the Industrial Revolution. According to Landes (1986), "...patents were not always the best way to protect knowledge. Instead, inventors preferred to try and keep devices and techniques secret, sometimes by so dividing the process that no one worker could penetrate the technique. This is what the great watchmaker Abraham-Louis Breguet proposed to do when he planned the mass production of watches by means of power tools and interchangeable parts: the aim was … security." (Landes, p. 615.)

Apple Inc. offers a modern case study in compartmentalism. In his book *Inside Apple* (2012), Adam Lashinsky explains that compartmentalism at Apple is at least in part physical:

> Apple employees know something big is afoot when the carpenters appear in their office building. New walls are quickly erected. Doors are added and new security protocols put into place. Windows that once were transparent are now frosted. Other rooms have no windows at all. These are called lockdown rooms: No information goes in or out without a reason. (p. 31)

In addition, Apple has in place an elaborate security clearance system to control information flow between and even within teams. "To discuss a topic at a meeting, one must be sure everyone in the room is "disclosed" on the topic, meaning they have been made privy to certain secrets....As a result, Apple employees and their projects are pieces of a puzzle. The snapshot of the completed puzzle is known only at the highest reaches of the organization." (p. 41)

To test our model of team formation and persistence in industrial R&D, we exploit a panel data set of researchers that includes all inventors listed on U.S. patents since 1975. Because a patent lists each inventor who instrumentally contributed to the development of the underlying invention, we are able to construct measures of teams in industrial R&D. We show that teams have been increasing in size and that the number and impact of lone researchers are falling. We show teams are remaining together over longer periods of time and more projects. The results from our regression analyses of team size and team persistence suggest that scientist mobility as well as firm-level technological characteristics and coordination cost do matter for both. To identify the effect of mobility on team size and persistence we take advantage of state-year variation in non-compete enforcement including a policy 'experiment' in the 1980s. In 1986, the state of Michigan reversed its long-standing opposition to non-compete covenants in employer contracts (Marx, 2009). A non-compete covenant is a promise by a worker not to work with a direct competitor for a fixed period of time following the end of employment. Because non-compete covenants are commonly incorporated into employment agreements of researcher employees, we argue that in states and years where non-compete agreements are enforced, the risk that the researcher will appropriate valuable intellectual property is lower.

The paper is structured as follows. Section 2 reviews recent empirical work describing teamwork in science and in industrial R&D. Section 3 gives a brief overview of our dynamic model of team formation. We describe how we will test its implications for team size and persistence and how these quantities should move with changes in the labor market and the innovative environment. In section 4 we describe our empirical approach and our data. Section 5 reports our empirical results. Section 6 discusses and concludes.

2. Literature Review

A number of studies report an increase in the size of teams in science and R&D. Adams, Black, Clemmons, and Stephan (2005) document increasing team size in science, where team size is measured by the number of authors on a scientific paper. Wuchty, Jones, and Uzzi (2007) looking at authorship counts on scientific papers and inventor counts on patents find that the size of teams has steadily increased since the 1960s. In addition, as assessed by the numbers of citations patents and papers received, the impact of works with multiple authors/inventors is greater than works with a single author/inventor. These studies show that these phenomena are exhibited widely across technology and scientific fields.

A number of studies examine the determinants of teamwork and collaboration. Arora and Gambardella (1990) offer evidence that collaboration arises from complementarities in skills. Mairesse and Turner (2005) investigating a sample of French physicists find that distance matters in whether the physicists collaborate. Agrawal and Goldfarb (2008) present evidence that the rise of the internet had contributed to increased collaboration. Rosenblat and Mobius (2004) also examine the importance of communication technologies on collaboration and networking.

A literature investigates the impact of team composition on team productivity, Hansen et al. (2006) use classroom performance on group projects to examine whether gender/racial diversity affects performance. Skilton (2009) examines whether team members' human capital affects productivity (citations) of teams.

The management literature examines persistence in collaboration (Skilton and Dooley, 2010, and Guimera, Uzzi, Spiro, and Amaral, 2005). Whether a team has worked together before may affect its productivity, for example, if repeated interaction increases trust or improves coordination or communication. Akgun and Lynn (2002) find that in product development teams in R&D-performing firms a team's longevity has a positive effect on productivity-related outcomes including team learning and cycle time, but not when there is a high degree of market and technical turbulence. Katz (1982) reports that the R&D team’s longevity's effect on productivity is possibly quadratic, peaking at two to four years from the team's inception. We are unaware of any analysis of this aspect of teamwork in the economics literature.

There exists a theoretical literature in economics on teams. Marschak and Radner (1972) and Cremer (1980) discuss the coordination of tasks in teams while Holmstrom (1982) explores incentives in teams. Dessein and Santos (2003) and Corts (2007) argue that teamwork can be used to solve multitask problems and suggest that increased use of teamwork in the manufacturing and service sectors may be due to improvements in performance monitoring which reduce the inefficiency associated with joint accountability. Alchian and Demsetz (1972) argue team production is more efficient within firms because firms lower transaction costs (e.g., monitoring). Other work explores optimal team creation as trading off gains to specialization and coordination or communication costs (Becker and Murphy, 1992; Bolton and Dewatripont 1994). In the context of science and innovation, Jones (2005) argues that accumulating knowledge increases the educational burden of scientists and inventors causing them to specialize which leaves them less able to innovate on their own.

In this paper we consider a dynamic model of team formation within R&D firms that incorporates the trade-off between specialization gains and coordination costs. The model emphasizes however the risk of researcher appropriation and allows team formation to be an element of an IP protection strategy. Worrying that research employees will steal intellectual property, an innovating employer may either reduce the number of researchers involved to reduce the number of potential leavers or compartmentalize research teams with more researchers.[7] This model adds to the standard economic theory of teams the idea of compartmentalization and also allows for the team's productivity to be related to its longevity and to turbulence in the technological or market environment (Akgun and Lynn).

3. Summary of the Teamwork Model

Our empirical framework is derived from our model which we briefly outline here, along with its major empirical implications that we will test. (The model is fully described in appendix A.)

Consider a firm with multiple R&D projects. For each project, the firm must hire a team of researchers. Each project is defined by a (non-overlapping) interval of tasks of length R. All projects are symmetric so for exposition we will focus on a representative project.

A project cycle consists of two periods. In the first period, the researchers hired

[7] A “divide and conquer” response to the threat of worker appropriation also underlies the firm’s decision of hierarchical form in Rajan and Zingales (2001). They study how an entrepreneur with a unique critical resource (an idea, good customer relationships, or superior management technique) decides on the firm's structure in the face of the threat of employee expropriation of the resource. They show that if the threat of expropriation is severe enough, the entrepreneur adopts a horizontal organizational form, which corresponds to our compartmentalization story. If the threat is not severe, she adopts a vertical hierarchical structure. In addition, the size of the organization (analogous to team, here) will be smaller, the higher the possibility of expropriation

onto the project team develop a prototype. They also develop team-specific human capital and learn valuable intellectual property (IP) that they may exploit for the firm's competitors if they separate from the firm in the second or marketing period of the two-period cycle. The external value of the IP, which we denote as ρ, is random and not revealed until the second period. Rosenberg (1996) argues that a new technology originally invented for one narrow purpose often finds high or even higher value in another purpose in an entirely different industry and these out-of-context applications are unpredictable.[8] ρ captures Rosenberg's out-of-context kind of external value as well as the in-context kind such as when the firm, say, solves a technical problem that also vexes its competitors. The team-specific human capital benefits the firm by reducing the firm's coordination costs if the team remains together for projects in subsequent periods.

In the second period, the firm produces and markets a product based on the prototype. The product's life ends at the end of the second period. Also in the second period, the firm begins a second two-period cycle on a new round of projects, either with the researchers from the first period or if these researchers separate from the firm a new set of researchers. As in the first round, researchers develop new prototypes (and learn IP) that leads to new products which are produced and sold in the next stage of the cycle, the third period. The life-span of the firm is three periods. In the third period,

[8] Rosenberg (p. 104) relates this story of the dawn of the computer.

> Howard Aiken, a Harvard physics instructor who was a great pioneer in the early development of the computer, continued to think of it in the narrow context in which its early development took place--that is, purely as a device for solving esoteric scientific problems. As late as 1956 he stated: "if it should ever turn out that the basic logics of a machine designed for the numerical solution of differential equations coincide with the logics of a machine intended to make bills for a department store, I would regard this as the most amazing coincidence that I have ever encountered" (Ceruzzi 1987, p. 197).

because R&D ceases all researchers depart the firm, using the IP they learned in the second period to compete against their former employer.

The production technology follows Becker and Murphy (1992). For each project output (revenue), Y, is Leontief and equal to $\min_{0 \le t \le R} Y(t)$, where Y(t) is the revenue from task t. The firm hires researchers each of whom is specialized in a subset of tasks in R. The output on a task depends crucially on the number of persons in a team. Define n as the number of persons hired onto a team in the first period. The more persons on a team the greater the degree of specialization and also the more time each worker can spend on a task. Thus, Y increases with n. However, larger teams impose a greater wage bill and entail greater coordination costs for the firm.

Both the researchers and the firm are risk neutral. The firm maximizes its discounted expected profits and each researcher maximizes his discounted expected wages. In the first period, the researcher will choose to work on the firm's team if his discounted expected earnings stream is higher with the firm than in alternative employment. Bargaining is such that the researcher's wages in the first period are fully discounted by the expected external value of the IP.

In the first period, the firm chooses team size to maximize its discounted profit. Thus, the firm takes into consideration the researchers' wages, the productivity advantages of specialization, and the costs of coordination. In addition, the firm's team size decision reflects future considerations. For example, by assembling a larger team the firm can ensure that each researcher operates only on a narrow set of tasks so in the second period if a researcher eventually leaves for a competitor she is less able to recreate the project on her own. This model assumes that only when researchers separate can they compete against the firm and also that separation guarantees competition. Compartmentalization reduces both the future likelihood that researchers separate and the costs their competition imposes on the firm if they do separate. But by reducing the breadth of the researchers' tasks and thus of the IP the researcher learns, the firm raises the wages it must pay its researchers in the first period. As in Pakes and Nitzan (1989), ρ is a kind of general human capital with the same implications for wages.

At the beginning of the second period two types of uncertainty are resolved which determine whether the research team hired in the first period remains intact or disbands. First ρ is realized. Second the firm learns the extent that technology or the market shifts for the next round of research projects. Researchers from the first period leave if their gains from leaving and competing exceed their employer's losses. Thus for example the greater the realized ρ, the more likely the separation condition will be met. The bigger the impact of experience on coordination costs, the less willing the firm is to see its first period researchers leave. Changes to technology or the firms' market alter the optimal sorting of tasks and, because researchers are specialized to tasks, of researchers into teams. Thus if the firm wishes to keep its researchers from the first period and to maximize its revenue from second period projects, it should regroup its researchers. But regrouping researchers destroys at least some of the team-specific human capital built up in the first period, reducing the firms' desire to keep the first period researchers through the second period and increasing the likelihood that the separation condition is met.

The model implies a number of hypotheses we will test which we now describe.

Team Size Hypotheses

A rightward shift in the distribution of the IP's external value, ρ, can have a positive effect on team size through a couple of channels. Corresponding to the effect of general training in labor economics, improved outside opportunities lower the first period wage, making larger teams less expensive for the firm. (This effect dominates an opposing effect of rising wages in the second period to retain workers due to their improving outside opportunities). In addition, a rightward shift in the distribution of ρ increases the benefit of additional workers, as reducing the knowledge each researcher acquires (compartmentalization) becomes more valuable.

Through other channels, however, a rightward shift in the distribution of ρ may lower n. Because firms anticipate more researcher departures, an increase in the value of external opportunities raises their expected costs of R&D. This scale effect reduces team size. Additionally, suppose the optimal first period team size, n, is smaller than the team size that maximizes the profit in the second-period project so that the marginal profit of n is strictly positive in terms of the profit from the second-period project. If the probability that the researchers stay falls due a rightward shift in ρ, the expected marginal profit from the second-period project when researchers stay falls and thus the optimal n falls. The last effect, which we will call the size trend effect, would make the optimal n increased or intact if the optimal n is bigger than or equal to the optimal team size in the second period, respectively, when ρ shifts rightward.

Hypothesis 1: Team Size-Mobility The effect of a shift in the distribution of ρ has an ambiguous effect on *n*. Higher anticipated mobility will increase team size if the compartmentalization plus the wage effects dominate, and will lower team size if the scale effect plus the negative size trend effect dominate.

Most other factors that we wish to test affect mobility, which has an ambiguous effect on team size. Our aim is to generate testable empirical implications, thus in our empirical examination of the determinants of team size we will control for mobility. The hypotheses below assume mobility is held constant.

Decreasing coordination costs or increasing researcher's productivity makes larger teams more attractive to the firm. This leads to the following two hypotheses.

Hypothesis 2: Team Size-Coordination Cost Higher coordination costs lower n, holding mobility constant.

Hypothesis 3: Team Size-Researcher Productivity Higher researcher productivity within the team raises n, holding mobility constant.

Because higher *general* productivity increases the researcher’s reservation wage it increases the cost of larger teams.

Hypothesis 4: Team Size-General Productivity Higher researcher reservation wages lower n, holding mobility constant.

When the technologies in the second period is different from the first period technology so that the second period revenue-maximizing grouping of researchers differ from the first period one, the second period productivity of researchers falls, which decreases the team size. But also researchers are more likely to move in this case and the added mobility can then increase or decrease team size. Thus after controlling for mobility in our empirical specification, we expect to find an adverse effect of technological change on n.

Hypothesis 5: Team Size-Technological Change Higher anticipated shifts in technology lower n, holding mobility constant.

The effect of task range, R, on team size is ambiguous. On the one hand, the individual researcher's breadth of knowledge increases with R, which gives the researcher more to appropriate if she leaves the firm, lowering her first period wage. This breadth effect increases n. On the other hand, a larger R lowers the gains to specialization in the first and second periods, reducing n.

Hypothesis 6: Team Size-Task Range Higher R has an ambiguous effect on n, holding mobility constant. If the breadth effect dominates, n will increase. If instead the specialization effect dominates, n will decrease.

Team Persistence Hypotheses

A rightward shift in the distribution of ρ produces two opposing effects on the persistence of teamwork. Mobility directly reduces teamwork persistence because researchers are less likely to work together again when they depart the firm. On the other hand, if team size increases as the firm compartmentalizes its R&D, the firm will have more accumulated team-specific human capital to exploit by retaining its researchers from the first period. This will raise teamwork persistence. However, with team size held constant, we will only have the former, negative effect. Because in our empirical work we are able to control for team size, we will ignore a parameter's effects on team persistence through team size in the hypotheses concerning ρ and the other parameters.

Hypothesis 7: Team Persistence-Mobility Higher ρ decreases team persistence, holding n constant.

Coordination cost, researcher productivity, and researcher reservation wage can affect persistence only through the scientist mobility channel, and their effects depend on whether n is bigger than the optimal second period team size or not. Consistent with the observed increase in team size, the following hypotheses assume that the optimal n is smaller than the optimal second period team size. An increase in the reservation wage will make separation less likely because, with increasing team size, the total wage expense to the firm is increasing. Similarly, higher coordination costs and lower researcher productivity will lead to greater persistence.

Hypothesis 8: Team Persistence-General Productivity Higher reservation wage increases team persistence, holding n constant.

Hypothesis 9: Team Persistence-Coordination Cost Higher coordination costs increase team persistence, holding n constant.

Hypothesis 10: Team Persistence-Researcher Productivity Higher researcher productivity decreases team persistence, holding n constant.

When technology is changing rapidly, the firm's profit is reduced in the case where researchers stay. This will increase mobility and reduce the persistence of teamwork. Furthermore, the optimal sorting of tasks into projects will deviate greatly from the optimal sorting of the first period (less overlap), which will reduce the persistence of teamwork.

Hypothesis 11: Team Persistence-Technological Change Greater technological change reduces team persistence, holding n constant.

Finally, an increase in project range, R, will increase the breadth of knowledge each researcher can acquire, which increases the value of their outside opportunities and hence mobility. But the improved outside opportunities lower first period wages causing n to rise, and firms are more likely to choose projects that are similar to those in the first period to take advantage of accumulated team-specific human capital, increasing persistence. Holding n constant, we would thus predict that an increase in R would reduce persistence.

Hypothesis 12: Team Persistence-Task Range The task range R has a negative effect on team persistence, holding n constant.

4. Data description

4.1 Descriptive Statistics

We use US patent data for our empirical analysis, equating the inventor lists on patents to inventor teams. Patents legally must name as inventors all persons who conceived a portion of any claim made by the patent.[9] The standard for co-inventorship does not require that each inventor contributed to the conception of all claims, or that co-inventors physically worked together (though some demonstration of collaboration and connection among the inventors is required). Patent law narrowly circumscribes co-inventorship. For example, laboratory directorship is sufficient in many disciplines to earn a scientist co-authorship on a publication, but it is insufficient to obtain co-inventorship. Also, persons who are technicians or any others who merely took direction from an inventor do not legally qualify as co-inventor. Contributing ideas,

[9] See Manual of Patent Examining Procedure (Eighth Edition, August 2001, revised July 2008), Chapter 2137.01, Inventorship [R-3] available online at http://www.uspto.gov/web/offices/pac/mpep/documents/2100_2137_01.htm (accessed June 25, 2009).

suggestions, or materials, even if the help proved crucial in bringing about the invention, is also insufficient for co-inventorship. The litmus test is in fact whether the person contributed to the conception of a claim. While in practice the conception test may not always be followed, including persons on the patent application who do not meet this test, or excluding persons who do, risks having the patent invalidated (see Crawford and Kokjohn, 2009, and Remus and Personick, 1995).

For our analysis of team persistence, we must be able to follow inventors from patent to patent. The patent data, however, identify inventors by name only, and do not provide unique identifiers for inventors. We use the inventor "disambiguation" produced by Lai, D'Amour, Yu, Sun, Torvik and Fleming (2011) as our source of inventor-level panel data.[10]

Team Size:

Figure 1a describes the average inventor team size by year. This figure includes only patents assigned to U.S. companies and corporations that are ultimately granted, by year of application. One can see that the average team size has steadily increased from 1975 through the early 2000s—by about 1 inventor over the period or by 62 percent. We find that the number of extremely large inventor teams (upwards of 20 or more inventors) has also increased (not reported). In contrast, the fraction of patents with a single inventor declined steadily during this period from the majority of patents (58 percent) in 1975 to a minority (33 percent) in 2003 (see Figure 1b).

The average team size by year by patent technological category is reported in Figure 2. Figure 2 shows that the increase in team size occurs across all of the

[10] We use the disambiguation produced by the "UPPER" parameterization of their algorithm.

categories, but that the increase is greatest in Chemicals and Drugs & Medical categories. In Figures 1a and 2, we note a distinctive blip in 1995. Prior to 1995, US patent protection ended 17 years after the patent *issuance* date. To comply with the Uruguay Round Agreements of the General Agreement on Tariffs and Trade (GATT), the US in 1995 extended patent protection to 20 years after the patent *application* date. A provision of the new law designed to ease the transition guaranteed that patent applications filed before June 7 and issued after June 8 enjoyed a monopoly period equal to 17 years post issuance or 20 years post application, whichever was longer. The blip may reflect a rush to file applications before June 7, 1995 to take advantage of the extended monopoly period. If patents with a longer shelf life are more valuable, and more valuable patents are produced by bigger inventor teams, an increase in team size in that year would seem natural. Consistent with the fact that drug patents have longer shelf life, the blip is most pronounced in the Drugs & Medical category.

Figure 3 describes the trend in team size by country or region of the patent's first inventor. Japan's inventor teams are the largest, followed by those of Europe and the UK. Between the 1970s and the early 2000s, Japanese teams have remained large and relatively constant in average size which ranged between 2.5 and 2.8 inventors per patent. By the end of this span, the European and North American mean inventor team sizes had nearly caught up to the Japanese. The fact that Japanese teams have remained constant in size indicates that rising team size in innovating US companies may not be exclusively due to changes in technology toward more specialization since Japan likely experienced similar technological changes as in the US.

Team Persistence:

We use the occurrence of multiple patents featuring the same inventors in a given window of time as an indicator of persistence. Teams make themselves visible when they patent and these patents reveal teams' size and composition–the number and identity of scientists working together. If the same subsets of inventors appear on multiple patents over time we know the teams persist. If patenting rates remain constant, then a rise in "repeats" means that persistence is increasing. If persistence is increasing, it means either that the projects in which teams are working are lasting longer, or that the teams are serially working on more projects.

We first examine the repetition of the same pairs of inventors. To measure persistence in year T, we ask what fraction of inventor pairs formed in year T, form again within three years. A possible repeating pair is identified in the following way: inventors A and B are on a patent applied for in year T.[11] If inventors A and B are found on a patent application dated within three years of their year T appearance, A and B are considered a repeating pair. (Note that the date of *first appearance* may not be the first patent that A and B are on together.) Table 1 shows the fraction of pairs that repeat by year. Table 1 shows persistence rising: 28% of pairs that appear in 1975 repeat sometime in the subsequent three years. By 1995, the fraction reaches 40%.

Some of what we call persistence in the team may be due to the following: a research team works on a project and the project produces multiple patentable outputs simultaneously. These patentable outputs result in a cluster of patents which cause pairs to be recorded as persisting. Because of the difficulty linking patents to R&D projects, our measure of persistence captures both the notion of the inventor pair

[11] The patent may have more than two inventors. The possible pairs from a patent containing inventors A, B, and C, say, are A-B, A-C, and B-C.

staying together from project to project but also the inventor pair coming together for only one project that produces multiple patents. Because patents that are filed three or more months apart are more likely to be from different projects, we can formulate a measure of project-to-project persistence by counting only those patents that are filed at least three months after the original pairing. Indeed, the persistence measure falls by about one fifth when we omit the first three months of the span (far right column); however, the rise in persistence over time remains. This makes us more confident that teams indeed are increasingly staying together for multiple projects.[12]

Table 2 shows the result from an analysis of inventor trios. Though trios repeat slightly less often than pairs, they too show an upward trend in persistence.

Table 3 conveys a sense of the duration of teams. It shows for pairs that are observed on the same patent in a year, the fraction that are observed on patents at various intervals subsequently. For example, of the inventor pairs observed in 1975, 10% pair again on patents applied within 3 months of their 1975 pairing, 12% on patents that are applied for between 3 months and 12 months following their 1975 pairing, etc. In the third year following their 1975 pairing, a pair's likelihood of appearing on a patent together is 6%. Table 3 shows that the incidence of re-pairing of any of the pair samples falls the further one gets from the date the pairs were identified. So, for example, for the 1995 pairs, the fraction that are again observed on a patent falls from 18% in the second year following their 1995 patent, to 14% in the third year and

[12] Until the late 1980's the number of patents granted had remained fairly steady, but since then the number has soared. Increasing patent rates may bias team persistence upward in recent years because we can identify repetitions of pairs only when patents are issued. However, because rising persistence has started well before 1990 the upward trend in team persistence we argue is not due to the rising number of patents.

8% in the fourth year following their 1995 patent. This pattern is observed for all five inventor samples. Note that almost half of the re-pairing that occurs in the first year following the base year pairing occurs in the first 3 months. For the 1995 pairs, most of what re-pairing occurs in the first year occurs in the first 3 months.

Table 3 shows that the duration of collaboration between inventors is increasing. If we observe a pairing of inventors in 1975, their probability of pairing again sometime between two and three years later is about .09. For pairs observed in 1995, that likelihood is greater—about .14. For pairs observed in 1975, their likelihood of pairing sometime between three and four years later is .06. For pairs observed in 1995, that likelihood is .08.

4.2 Model Specification of the Regression Analyses

This section describes our specifications of the regression-based tests of the hypotheses enumerated in Section 3. In our analysis of team size, the unit of observation is a patent and the dependent variable is the number of inventors appearing on a patent. The analysis includes all patents applied for between 1975 and 2004.[13] In the team persistence analysis, in the first set of regressions each observation is a pair of inventors that worked together on a patent that has two or more inventors. For patents that had more than two inventors, we randomly chose one pair (e.g., four inventors on a patent imply six pairs of which we randomly choose one to include in the analysis). The dependent variable is the number of additional times between three months and three years of first appearing that the pair reappears on a patent, and the analysis includes all inventor pairs that occur on patents applied for between 1975 and 2001.

[13] Although the data from Lai, D'Amour, Yu, Sun, Torvik and Fleming (2011) include patents filed in 2005 and 2006, we excluded them in our analysis due to the truncation problem.

The basic specification for our regression analysis of team size and persistence is a Poisson model with firm-level fixed effects. We employ the Poisson regression method because in both sets of analyses our dependent variable is a count variable with nonnegative values. Our models also include dummy variables for six patent technological categories (Chemical, Computers & Communications, Drugs & Medical, Electrical & Electronic, Mechanical, and Others) and filing year dummies.

The basic set of explanatory variables includes a dummy variable for states' enforcement of non-compete covenants, the number of inventors within 50 miles of the first inventor on the patent, firm-level R&D expenditures, industry-level R&D expenditures, and the median wage for scientists by industry. In the team persistence analysis, these variables are measured in the first year the pair appears in our data. We also use as additional regressors the average number of patents per inventor in the previous three-year period, average distance between inventors, a measure of generality of a patent, the number of citations received in the first five years following the granting of the patent, and the number of claims. These additional regressors are not included in our basic set because they may be endogenous. All regressors except those which can take a zero value are in logarithmic form.

We do not directly observe ρ, the instigator of mobility in our model. Instead we use variation in enforcement of non-compete clauses to investigate the relationship between scientist mobility and team size or persistence. A non-compete covenant is a promise by a worker not to work with a direct competitor for a fixed period of time following the end of employment. Non-compete covenants are commonly incorporated into employment agreements of senior research employees. Many states enforce non-compete covenants, but some are reluctant to enforce them because of the restrictions they place on the worker's ability to secure employment (see Dworkin and Callahan, 1998; Gilson, 1999; Koh, 1998). Previous research appears to show that states' enforcement policies are of economic consequence. Fallick, Fleischman and Rebitzer (2006) find evidence that non-compete enforcement reduces firm-to-firm mobility of computer workers. Using Michigan's 1985 reversal on its long-held refusal to enforce non-compete covenants, Marx, Strumsky and Fleming (2009) show enforcement of non-compete covenants reduces the mobility of certain kinds of inventors. Stuart and Sorenson (2003) find that start-ups are more frequent in regions where non-compete clauses are not enforced. We use variations in enforcement of non-compete covenants across states over time to evaluate the importance of worker mobility in limiting teamwork. Following Marx et al. (2009) and Marx (2009), we use Malsberger (1996) to identify the states that restrict non-compete enforcement. According to Malsberger, the following states had specific legislation restricting enforcement of non-competes: Alaska, California, Connecticut, Minnesota, Montana, North Dakota, Nevada, Oklahoma, Washington, and West Virginia. In addition, prior to March 1985, Michigan restricted enforcement of such contracts.

We use the number of inventors within 50 miles of the address of the first inventor as a proxy for the outside opportunities available to the inventors on the patent. Isolated inventors will have higher costs of moving and will show lower mobility rates. The lower mobility may lead to smaller or larger teams depending on whether the scale or compartmentalization effect dominates. On the other hand, denser labor markets permit greater specialization and therefore larger teams (Smith, 1976; Becker and

Murphy, 1992). In sum, teams may be larger or smaller in areas with fewer inventors depending on whether the scale effect dominates the specialization and compartmentalization effects.

We obtain firm-level R&D expenditures from the Compustat database, using the link between the patent assignees and the Compustat firms created by the NBER Patent Data Project (PDP) (Iain Cockburn, PI). Our analyses are based on patents applied for between 1975 and 2004. We use the firm's R&D expenditures as a measure of the size of its research enterprise. For firm-years where R&D is reported as zero we substitute in the minimum non-zero R&D expenditure in the data and include an indicator that is equal to one for such firm-years (and equal to zero otherwise). This allows us to take the natural log of R&D without dropping firm-years from the analysis. Besides the scale effect in research, this variable may reflect the maturity of the firm's technology.

We use as a measure of maturity of the industry where the firm belongs, the log of industry productivity, defined as sales net of (non-labor) costs per worker. And firms in a mature industry may face slower technological change. The industry classification for this measure is roughly the same as the two-digit code for classifying industries used by Bound et al. (1984).

Our proxy for the productivity of scientists employed by the firm, or reservation wage, is the median wage and salary for scientists in the firm's industry.[14] The data for scientist wage are taken from the Annual Demographic Files (March Supplements) of

[14] We include the following occupation categories for scientists and engineers (the three-digit 1980 standard occupational classifications are in parentheses): Engineers (044-059), Mathematical and computer scientists (064-068), Natural scientists (069-083), Clinical laboratory technologists and technicians (203), Engineering and related technologists and technicians (213-216), Science technicians (223-225), and Computer programmers (229).

the Current Population Survey (CPS), conducted by the U.S. Census Bureau. The median wage is calculated from annual wage and salary of all scientists and engineers. The March CPS yields records on 2,600 scientists and engineers on average annually between 1975 and 2004. This variable as well as the firm-level R&D and the industry productivity is in constant 2009 dollars.

We also include the number of claims, a measure of generality of a patent, and the number of citations received in the next 5 years. The measure of generality is an index which ranges between zero and one, and is higher for patents that are cited by subsequent patents from a wide range of fields. If we consider forward citations as a measure of the impact of a patent, a high generality value suggests that the patent has a widespread impact (Hall et al., 2001). We thus regard the generality measure as a proxy for the project rage (R) in our model. The number of claims can be used as a measure of the scope or the width of the invention (Lanjouw and Schankerman, 2004) and may also proxy the project range (R) in our model.

We include the number of citations received as a regressor to account for the success of the team in its first observed encounter. Alternatively, this variable can be considered as a proxy for the researcher's productivity in the firm.

The data for both the average number of patents by an inventor in the last three years and the average distance between inventors are taken from the inventor-level patent data constructed by Lai et al. The former variable is used as a measure of an inventor's research productivity, or as a control for patenting rate in the team persistence analysis. The latter variable reflects variations in coordination cost across firms: a firm with lower coordination costs is likely to show a shorter distance between its inventors.

Here we digress briefly on how we align in time patent, firm and other data. Because of the examination process, a patent is often granted several years after the patent is filed. The filing or application date listed in the USPTO's records is often an adequate approximation of the date when the invention was completed, and thus an adequate approximation of the end date of the inventor team's work that lead to the underlying invention. Most researchers use the application date as a stand-in for the date of invention in empirical work involving patents. In instances where the application is a "child" application of a prior "parent" application—that is, where the patent is granted from a "continuing application"—the application date may be several years removed from the invention date.[15] We therefore use the earliest application date at the head of any string of continuing applications that lead to the patent as the invention date.[16] In other work, we have found that this date performs better than the application date.

[15]After an inventor files a patent application, s/he is entitled to file one or more continuing applications later claiming the priority date of the original or parent patent application. The ability to file continuing applications is valuable because it allows the inventor to modify the claims of the original application while still obtaining the benefit of the parent's priority date. To file a continuing application claiming the parent's priority date, a number of conditions must be met, including that the new claims must be fully supported by the subject matter disclosure in the parent application, and the parent application must not yet have been issued into a patent or abandoned (see 35 U.S.C. §119 and §120). Each continuing application can have one or more continuing applications filed on it which leads to potentially long chains of applications and many years of pendency for inventions that were complete as of the original parent's filing date. As an example, US patent no. 7,629,736 was granted in 2009 from an application that was filed in 2005, which was the culminating child application of a string of five continuing applications, traceable to an original patent application filed in 1994. For this patent granted in 2009, we use 1994 as the year of invention meaning that in our analysis we relate it to the assignee's expenditures, the legal environment, etc. as measured in 1994. For the years of our study, approximately one quarter of patents were granted from continuing applications.

[16] This is the patent's "priority date." If the earliest application date was preceded by an application in a foreign jurisdiction, we use instead the foreign application (priority) date as the invention date.

Table 4 displays the definitions and some simple statistics of the variables used in the regression analysis.

5. Results

5.1 Results from the Team Size Analysis

Table 5 shows the results from the analysis of team size. In the first column (Model 1) we report the results with only our basic set of variables as regressors. Note that in the Poisson specification the estimated coefficients for the log-transformed regressors have an elasticity interpretation.

In Model 1, the dummy variable for whether the first inventor on a patent is located in a state that enforces non-compete covenants is statistically significant and positive. This implies that all else equal, research teams that are in a state where non-compete clauses are enforced are larger by about 4 percent. If enforceable, non-compete covenants reduce the external value of knowledge transfer, and therefore inventor-employees in states where non-compete clauses are enforced are less likely to move. An interpretation of our finding is that a higher risk of worker appropriation results in net reduction in team size—that is, the scale effect dominates.

While the use of non-compete covenants in the employment contracts of key research personnel is now commonplace, we know of no evidence about how their use has changed over time. We therefore do not know how much of the increase in team size they might account for. In model 2 we interact the non-compete dummy with a time trend. The coefficient on the interaction term is insignificant suggesting that non-compete use has not increased over time.

The number of inventors within 50 miles of the first inventor on a patent has a positive and statistically significant effect on team size. This effect is, according to our interpretation, due to a greater supply of specialized workers or the KM model's compartmentalization effect.

We find consistently in Table 5 a negative (and statistically significant) relationship between the size of the firm's R&D enterprise and the size of its teams. We show that the median wage for researchers is positively correlated with team size, though the implied elasticity is small (about .03). This finding is consistent with the KM model where median wage is interpreted as reflecting researcher quality.

In Models 1 and 2, the estimated effect of industry productivity on team size is positive but insignificant. If this variable measures the maturity of an industry and a mature industry experiences slower technological change, the Team Size-Technological Change hypothesis predicts a positive relationship between industry productivity and team size.

Model 3 adds to the base specification the number of claims, the average number of patents produced by the inventors on the patent in the previous three years, the average distance between inventors' home residences, a citation-based measure of the generality of the patent, and the number of citations received by the patent in its first five years.

As a measure of an inventor's research productivity, we expect the average number of patents produced by an inventor in the last three years to have a positive effect on team size. However, our regression result shows the opposite effect. One possible reason for this is that persons who have had more patents in the past are older, and older inventors, either because they were more likely to have been trained as generalists (Jones, 2005) or because of habit, are less inclined to team with others.

Firms may take advantage of lower coordination costs—due to, say, improvements in communication technologies, cheaper air travel, and managerial innovations—by bringing in talent from farther away. We take average distance as an indirect measure of coordination costs and predict a positive association between distance and team size. This is confirmed in our regression result: the average distance between inventors is significantly and positively correlated with team size.

Table 5 shows that the generality measure, a proxy for the project range, is not significantly correlated with team size. The number of claims as a measure of the scope or the width of an invention—and possibly also capturing the project range—is shown to have a significantly positive association with team size, suggesting the breadth effect dominates the specialization effect.

As a measure of how successful a team is or its research productivity, we expect that the number of forward citations has a positive association with team size, which is confirmed in Table 5.

Table 5A reports the results of a difference-in-difference analysis that exploits the policy change in Michigan. All models use California-based patents as the comparison. California has not enforced noncompete covenants since its entry into the Union. The coefficient of interest is the coefficient on the interaction of Michigan and “After_1986”, indicators that the patent’s first inventor resides in Michigan and that the application was filed after the beginning of 1986, when Michigan first began enforcing noncompetes. We interpret that coefficient as identifying the effect of enforcing noncompetes on team size.

Model 4 includes all patents of all technology classes from the two states. Models 5 and 6 show the results when we compare similar patents in Michigan and California. Model 5 compares patents within the automobile and auto parts industry. Model 6 employs the Coarsened Exact Matching (CEM; Iacus et al 2011a and 2011b) procedure first on the sample before employing difference-in-differences. CEM matches to each Michigan patent a California patent that is identical by industry (according to 23 ARDSIC industry categories; see Hall and Vopel, 1997) and by groups of assignee size in sales, employment and R&D expenditures (where no matches can be found, the observation is omitted). The difference-in-difference analysis generates larger estimated effects of allowing noncompetes on team size. In model 4, the effect of noncompete enforcement increases team size by 12 percent; comparing only within the automobile/parts industries, allowing noncompetes raises team size by over 13 percent.

5.2 Results from the Team Persistence Analysis

The results of our team persistence analysis are shown in Tables 6. Table 6 features the same right-hand side specifications as in Table 5, with team size as an additional regressor. In Table 6 we investigate the persistence of inventor pairs.

As predicted by our model, holding team size constant, the dummy variable for non-compete covenant enforcement shows a positive effect on team persistence. Being in a state that enforces noncompetes increases the number of repeats the average inventor pair experiences in three years by at least 16 percent. Inventors in states that enforce non-compete covenants face lowered external value of knowledge transfer and are therefore less likely to move. This increases team persistence. Across specifications, firms with larger R&D enterprises show less persistence in their teams.

We find that the coefficient on log industry productivity is significantly positive. Because we control for team size, this result is consistent with our assumption that industry productivity reflects the maturity of an industry and is thus negatively related to the amount of technological change in the industry. Under this prediction, we would expect a positive relationship between industry R&D and persistence.

The median wage of scientists is shown in Table 6 to have a negative and significant effect on team persistence which contradicts hypothesis Team Persistence-General Productivity. One possible explanation for this finding is that the median wage reflects the productivity of researchers, and more productive workers tend to be older workers, and age reduces mobility (e.g., Hall, 1982).

Interestingly, the number of inventors in the area has a positive influence on persistence. We have controlled for team size, so the greater opportunities for moving represented by the larger local inventor populations should reduce persistence and yet it does not. Perhaps we are omitting a variable that is correlated with the density of inventiveness and team persistence.

Why might labor market size increase team persistence? The number of inventors on a patent (which is our team size variable in the regressions) will imperfectly reflect the size of research teams because for example a three-person team may have various pair combinations in different patents that they produce and our team size variable will be two although the real research team size is three. If larger labor markets induce larger research teams, and larger teams accumulate more team specific human capital, the market size variable will be positively correlated with persistence.

The results from model 3 show that the average number of patents of an inventor in the last three years has a significant and positive effect on team persistence. We think that this is mainly due to the fact that by construction teams with more patents are more likely to show up again in later patents just because they produce more patents, thereby showing higher persistence. As this variable is included as a regressor, we can control for a team's patent productivity and estimate the effects of other regressors without this concern.

A firm with lower coordination cost should have longer distance between its inventors. If the optimal team size is expected to grow over time, lower coordination cost reduces the incentive for firms to retain scientists and hence lowers persistence. This story is consistent with the finding in the tables: the average distance between inventors is negatively and significantly correlated with team persistence.

Table 6 shows that the number of claims is positively correlated with team persistence. If a patent with a larger scope and wider width is more valuable or important, inventors who produced it are a more successful pair and therefore are more likely to work together again.

A higher generality measure for a patent may imply that the underlying technology is more widely applicable in other fields or firms. If so, the inventors of such patents may be more frequently attracted away to other firms, which will reduce persistence.

As a measure of how successful a team is, we expect forward citations to have a positive association with team persistence. We find this expectation confirmed in the results reported in Table 6.

Table 6A reports the difference-in-difference results based on the Michigan policy change. We find, except for model 4, a stronger effect of allowing noncompetes on persistence.

6. Discussion and Conclusion

In this paper we show that earlier reported trends in team size extend to the mid-2000s. Between 1975 and 2003 mean inventor team size increased by 62 percent and the fraction of single inventor patents fell from 58 to 33 percent. We show that this trend extends across technological classes but the increase in team size has been greatest in Chemicals and Drugs & Medical categories. We also show that the trends in inventor team size by country of patent origin. Over the 1975-2003 period, the size of Japanese inventor teams on patents filed in the US have remained relatively steady and high by world standards, at about 2.5 inventors per patent. But by 2003 European and North American inventor teams had largely caught up with the Japanese. We also show evidence that inventor team size has responded to changes in laws governing patent protection during the mid-1990s.

We also show the first econometric evidence, we believe, that inventor teams are remaining intact for longer periods and over more projects. We find, for example, that two-person inventor teams formed in the mid-1970s had a 24 percent chance of appearing again on a patent within three years (omitting the first three months following their formation). By the mid-1990s that likelihood had increased to 33 percent. The analogous numbers for three-person inventor teams are 16 percent and 28 percent, respectively.

We then analyze the determinants of inventor team size. We use states' policies toward employment non-compete covenants to proxy for appropriation risk. We find that in states where non-competes are enforced, inventor teams tend to be larger. This finding is inconsistent with a compartmentalization story, but is consistent with firms scaling back teams to minimize appropriation cost. It is also consistent with a coordination cost story. In our model, coordination costs fall within a team as team members gain experience working with one another. If a firm anticipates its researchers will be retained, it will create larger teams to exploit the returns to specialization, because in an environment where researchers do not turn over, the firm is able to offset high initial coordination costs with lower coordination costs later on.

We also find some indirect evidence that coordination costs matter in the determination of team size. In firms and years where teams comprise inventors who reside greater distances from one another, team size tends to be larger. This is consistent with the following story: with the lowering of some kinds of coordination costs (such as communication-related costs), firms will respond by searching farther afield geographically to obtain the right kinds of inventor expertise. Firms will also take advantage of the fall in coordination costs to increase the size of the team. For this reason we argue that team size co-varying with average distance among the team's inventors is indirect evidence of a coordination cost effect on team size.

We analyze our model's implications for team persistence. We find that in states where non-competes are enforced, inventor teams are more likely to repeat. This is consistent with the appropriation part of our model: where non-competes are enforced, inventors will not leave an employer to appropriate the employer's IP elsewhere. Of course one does not need the appropriation story to explain an increase in persistence from non-competes. Enforcing non-competes make it harder for inventors to seek employment elsewhere for non-IP reasons as well. Other findings include: teams are more likely to persist when they are (1) located in smaller R&D shops; (2) on patents that describe less broad-based technological advances or on patents that assert more claims; and (3) on patents that generate more citations.

References

Adams, J.D., Black, G.C., Clemmons, J.R., and Stephan, P.E. (2005). “Scientific teams and Institutional Collaborations: Evidence from U.S. universities, 1981–1999,” Research Policy, 34, pp. 259-285.

Agrawal, Ajay and Goldfarb, Avi (2008) “Restructuring Research: Communication Costs and the Democratization of the University Innovation,” American Economic Review, 98(4), pp. 1578-1590.

Akgun, Ali E. and Lynn, Gary S. “Antecedents and consequences of team stability on new product development performance,” *Journal of Engineering and Technology Management,* Volume 19, Issues 3-4, September 2002, Pages 263-286.

Alchian and Demsetz. 1972. “Production, information costs, and economic organization.” American Economic Review 62 (December): 777-795.

Arora, Ashish and Alfonso Gambardella (1990) “Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology,” *The Journal of Industrial Economics,* Vol. 38, No.4., pp. 361-379.

Arrow, K.J. “Economic Welfare and the Allocation of Resources for Invention.” In R.R. Nelson, ed., *The Rate and Direction of Inventive Activity: Economic and Social Factors, NBER Special Conference Series*. Vol. 13, Princeton University Press, New Jersey, 1962.

Becker, Gary S and Kevin Murphy. 1992. "The Division of Labor, Coordination Costs, and Knowledge," The Quarterly Journal of Economics, MIT Press, MIT Press, vol. 107(4), pages 1137-60.

Bolton, Patrick & Dewatripont, Mathias, 1994. "The Firm as a Communication Network," The Quarterly Journal of Economics, MIT Press, MIT Press, vol. 109(4), pages 809-39.

Bound, J., Cummings, C., Griliches, Z., Hall, B.H. and Jaffe A. “Who Does R & D and Who Patents?” In Z. Griliches, ed., *R&D, Patents, and Productivity*, NBER, 1984.

Corts, K. S. 2007. “Teams versus individual accountability: solving multitask problems through job design.” The RAND Journal of Economics, 38: 467–479.

Cremer, Jacques. 1980. “A partial theory of the optimal organization of a bureaucracy,” Bell Journal of Economics, Autumn, 683-693.

Crawford, Bradley W. and Kokjohn, Sydney R. " "Navigating Inventorship in the Chemical Industry," *Snippets: Review of Developments in Intellectual Property Law*, Vol 7, Issue 1, Winter 2009, pp. 1, 5-7.

Dessein, Wouter, and Tano Santos. 2003. "The Demand for Coordination." NBER Working Papers 10056, National Bureau of Economic Research, Inc.

Dworkin, T.M. and Callahan, E.S. "Buying Silence." *American Business Law Journal*, Vol. 36 (1998), pp. 151-152.

Fallick , Bruce; Fleischman, Charles A.; and Rebitzer, James B. (August 2006) "Job-Hopping in Silicon Valley: Some Evidence of the Microfoundations of a High-Technology Cluster," *The Review of Economics and Statistics*, 88(3): 472–481

Gilson, R.J. "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not To Compete." *New York University Law Review*, Vol. 74 (1999), pp. 575-629.

Guimerà, R., Uzzi, B., Spiro, J., & Nunes Amaral, L. A. (2005) "Team assembly mechanisms determine collaboration network structure and team performance." *Science*, 308: 697–702.

Hall, Robert E., "The Importance of Lifetime Jobs in the U.S. Economy," *American Economic Review*, September 1982, 72, 716-24.

Hall, B. H., A. B. Jaffe, and M. Tratjenberg (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498.

Hansen , Zeynep; Owan, Hideo and Pan, Jie (May 2006) "The Impact of Group Diversity on Performance and Knowledge Spillover – An Experiment in a College Classroom," NBER Working Paper No. 12251

Holmstrom, B. "Moral Hazard in Teams." *Bell Journal of Economics*, Vol. 13 (1982), pp. 324-340.

Stefano M. Iacus, Gary King, and Giuseppe Porro, "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 2011, first published online August 23, 2011

Stefano M. Iacus, Gary King, Giuseppe Porro. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding," *Journal of the American Statistical Association*. March 1, 2011, 106(493): 345-361

Katz, Ralph. 1982. "The Effects of Group Longevity on Project Communication and Performance." Administrative Science Quarterly, Vol. 27, No. 1, pp. 81-104.

Kim, J., Lee, Sangjoon, and Marschke. G. "International Knowledge Flows: Evidence from an Inventor-Firm Matched Data Set," in *Science and Engineering Careers in the United States: An Analysis of Markets and Employment*, (Richard Freeman and Daniel Goroff, eds.), NBER and The University of Chicago Press, 2009, pp. 321-348.

Kim, Jinyoung, and Marschke, Gerald. "Labor Mobility of Scientists, Technological Diffusion, and the Firm's Patenting Decision," *The Rand Journal of Economics*, Summer 2005, 36(2), pp. 298-317

Koh, J.L. "From Hoops to Hard Drives: An Accession Law Approach to the Inevitable Appropriation of Trade Secrets." *The American University Law Review*, Vol. 48 (1998).

Ronald Lai; Alexander D'Amour; Amy Yu; Ye Sun; Lee Fleming, 2011, "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)",
http://hdl.handle.net/1902.1/15705 UNF:5:RqsI3LsQEYLHkkg5jG/jRg== V3 [Version]

Lanjouw, Jean O. and Schankerman, Mark. "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators," *The Economic Journal*, Vol. 114, No. 495 (Apr., 2004), pp. 441-465.

Lashinsky, Adam. 2012. *Inside Apple: how America's most admired-and secretive-company really works.* New York: Business Plus.

Mairesse, Jacques, and Laure Turner. 2005. "Measurement and Explanation of the Intensity of Co-publication in Scientific Research: An Analysis at the Laboratory Level." In *New Frontiers in the Economics of Innovation and New Technology: Essays in Honor of Paul David,* ed. Cristiano Antonelli, Dominique Foray, Bronwyn H. Hall, and W. Edward Steinmueller, 255–94. Cheltenham: Edward Elgar Publishing

Malsberger, B. M. (1996). *Covenants Not to Compete: A State-by-State Survey.* The Bureau of National Affairs, Inc.: Washington D.C.

Marschak, Jacob, and Roy Radner. 1972. *Economic Theory of Teams*. Yale University Press, New Haven and London.

Marx, Matt. (2009) "Good Work If You Can Get It (Again): Non-compete Agreements and Ex-employees." MIT Working Paper.

Marx, Matt; Strumsky, Deborah; and Fleming, Lee (June 2009) "Mobility, Skills, and the Michigan Non-Compete Experiment," *Management Science*, Vol. 55, No. 6, pp. 875–889

Pakes, Ariel and Nitzan, Shmuel. "Optimum Contracts for Research Personnel, Research Employment, and the Establishment of 'Rival' Enterprises," *Journal of Labor Economics* 1(4), 1983.

Rajan, Raghuram G. and Zingales, Luigi 2001 "The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms,, *Quarterly Journal of Economics*, 116 (3): 805-851.

Rosenberg, Nathan. Uncertainty and technological change. In R. Landau, R. Taylor, G. Write, eds. *The mosaic of economic growth*. Stanford University Press, Stanford, CA. 1996.

Rosenblat, Tanya S., and Markus M. Mobius. 2004. "Getting Closer or Drifting Apart?" *Quarterly Journal of Economics*, 119(3): 971–1009.

Skilton, Paul F. (2009) "Does the human capital of teams of natural science authors predict citation frequency?", *Scientometrics*, Vol. 78, No. 3 pp. 525–542

Skilton, Paul F. and Dooley (2010) "The Effects of Repeat Collaboration on Creative Abrasion," *Academy of Management Review*, Vol. 35, No. 1, 118–134.

Smith, Adam (1976), *An Inquiry into the Nature and Causes of the Wealth of Nations*, University of Chicago Press, Chicago.

Stephan, P. "The Economics of Science." *Journal of Economic Literature*, Vol. 34 (1996), pp.1199-1235.

Stewart, J. "Has Job Mobility Increased? Evidence from the Current Population Survey: 1975-1995." Bureau of Labor Statistics Working Paper No. 308, 1998.

Stuart, T., O. Sorenson. (2003). "Liquidity events and the geographic distribution of entrepreneurial activity" *Admin. Sci. Quart.* 48 175–201.

Wuchty, S., Jones, B.F., Uzzi, B. (2007). "The Increasing Dominance of Teams in Production of Knowledge," Science, 316, pp. 1036-1039.

Figure 1a

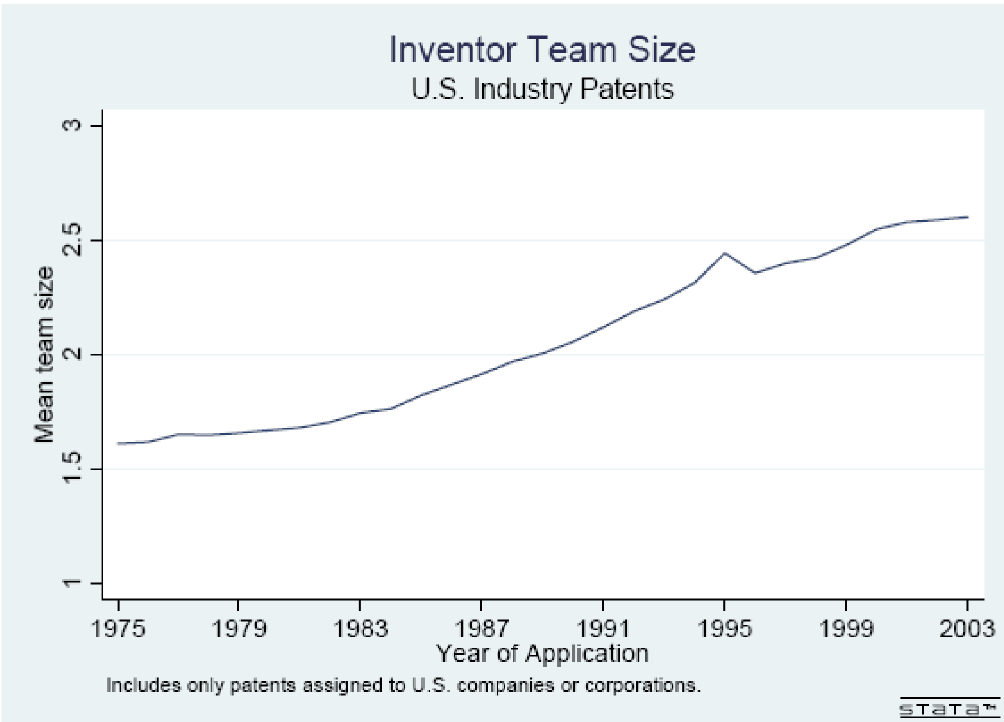

Figure 1b

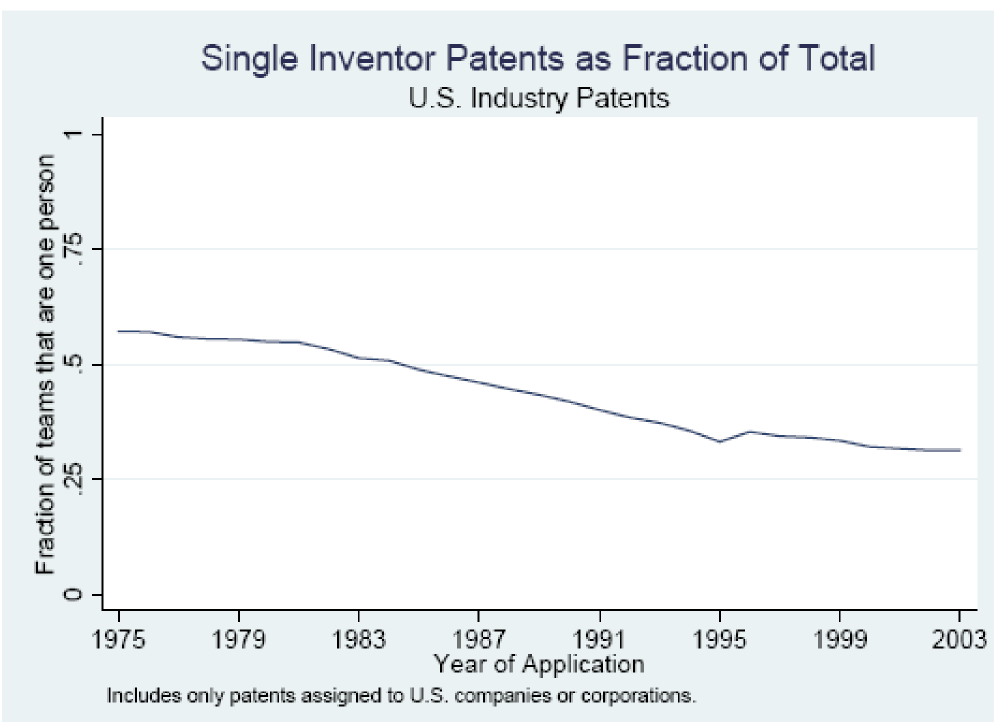

Figure 2

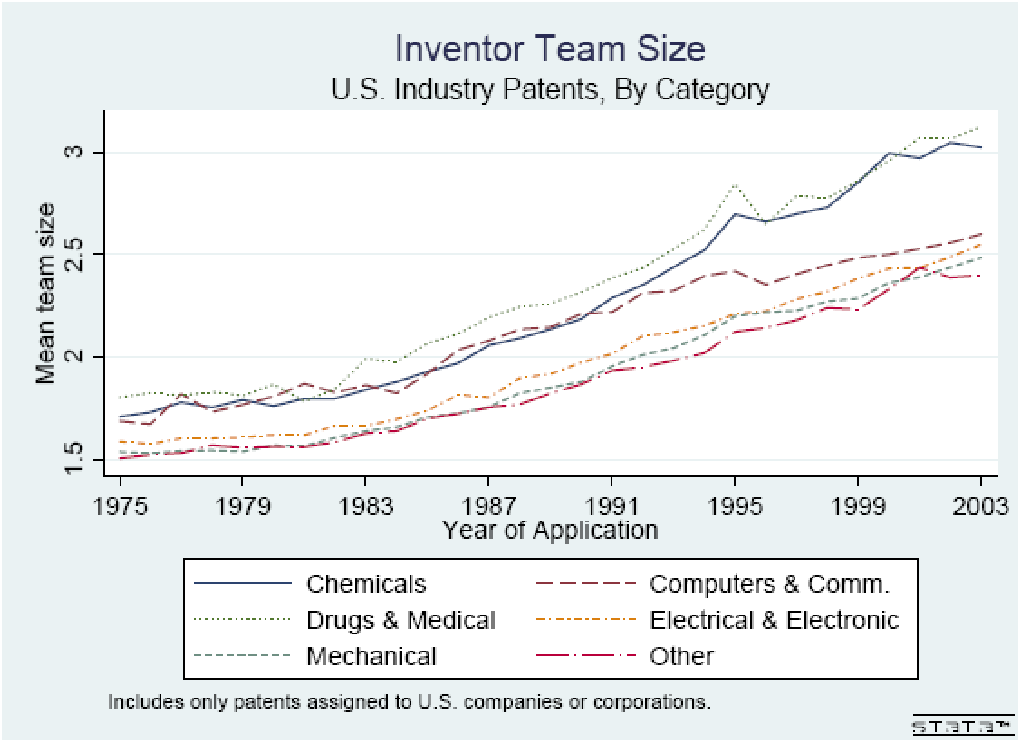

Figure 3

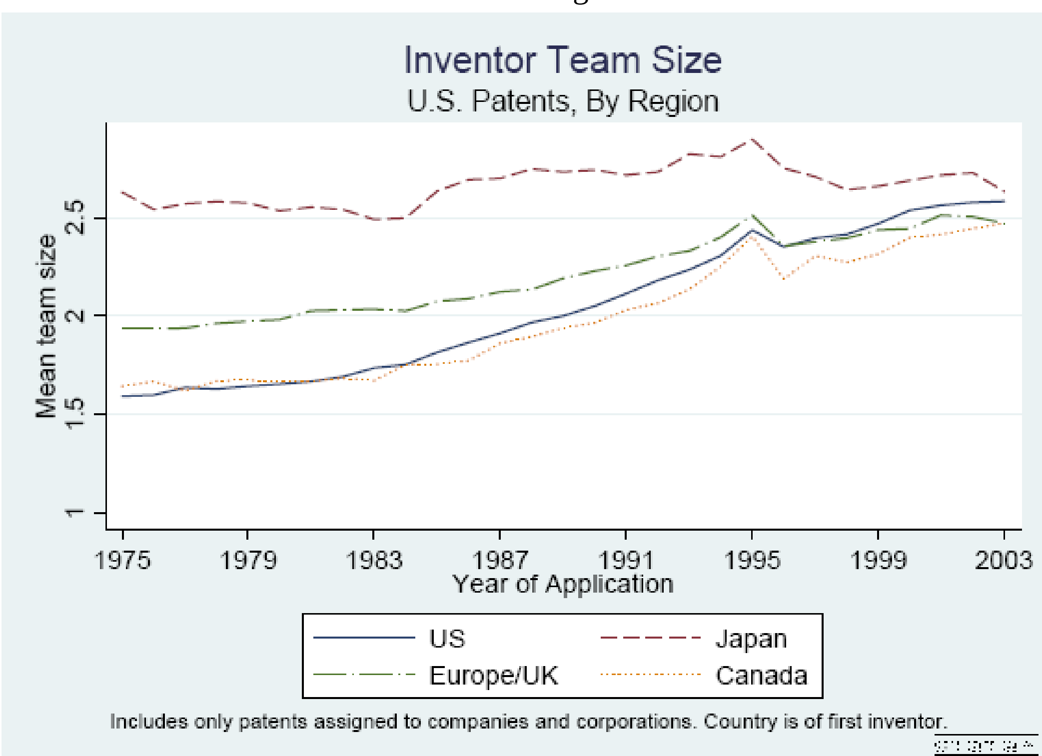

Table 1 Persistence of Inventor Pairs By Year		
	Fraction of inventor pairs that appear again sometime…	
Year of patent application date	0 to 3 years after patent application date	3 months to 3 years after patent application date
1975	.28	.24
1980	.29	.25
1985	.30	.26
1990	.33	.29
1995	.40	.33

Table 2 Persistence of Inventor Trios By Year		
	Fraction of inventor trios that appear again sometime…	
Year of patent application date	0 to 3 years after patent application date	3 months to 3 years after patent application date
1975	.19	.16
1980	.24	.19
1985	.22	.18
1990	.29	.25
1995	.37	.28

Table 3 Duration of Teamwork					
	Fraction of inventor pairs that appear again in the span				
	0 to 3 months after patent application date	3 months to 1 year after patent application date	1 to 2 years after patent application date	2 to 3 years after patent application date	3 to 4 years after patent application date
1975	.10	.12	.14	.09	.06
1980	.10	.12	.14	.09	.06
1985	.09	.12	.14	.11	.08
1990	.10	.12	.16	.12	.08
1995	.15	.13	.18	.14	.08

Table 4
Variable Description and Summary Statistics

Variable	Description	Mean	St. Dev	Min	Max
Team size	Number of inventors appearing on a patent	2.238	1.520	1	32
Number of repeats for pairs	Number of later patents for a pair of inventors to reappear together within 3 years of first appearance	1.176	4.805	0	148
Noncompete	=1 if the first inventor is located in a state that enforces non-compete covenants	0.688	0.463	0	1
Log number of inventors w/i 50 miles	Logarithm of the number of unique inventors residing within 50 miles from the first inventor's residence	7.370	1.193	0	10.11
Log R&D	Logarithm of R&D expenditures of a firm (constant 2009 GDP deflator)	1.100	1.895	-10.62	4.945
Log industry productivity	Logarithm of (sales-cost of goods sold + labor cost)/employment of a firm's industry (constant 2009 GDP deflator)	0.278	0.375	-1.846	1.660
Log median wage for scientists	Logarithm of median wage and salary of scientists by industry (constant 2009 GDP deflator)	6.397	0.202	5.747	7.225
Log number of claims	Logarithm of the number of claims on a patent	2.480	0.820	0	6.766
Average patents per inventor	Average number of patents per inventor over all inventors on a patent in the last 3 years	1.191	3.175	0	152
Average distance between inventors	Average distance between inventors on a patent (in miles)	250.8	795.1	0	11610
Generality	Index of how widely a patent is cited by subsequent patents in various technology fields	0.558	0.328	0	1
Citation received	Number of citations a patent received in 5 years following its granting	5.002	6.556	0	199

Table 5
Determinants of Team Size

Dependent Var.: team size — Poisson model

Variable	Model 1	Model 2	Model 3
Noncompete	0.03846 9.60	0.04711 7.50	0.03007 3.58
T*Noncompete		-0.00059 -1.79	
Log number of inventors within 50 miles	0.00309 3.17	0.00302 3.10	0.00451 2.14
Log R&D ($2009)	-0.00727 -3.44	-0.00759 -3.58	-0.01494 -3.24
Zero R&D dummy	0.12439 1.47	0.12273 1.45	0.09640 0.57
Log industry productivity ($2009)	0.00531 1.02	0.00574 1.10	0.02068 1.88
Log median wage for scientists ($2009)	0.03483 2.99	0.03441 2.95	-0.00383 -0.14
Log number of claims			0.04308 16.68
Average patents per inventor in last 3 years			-0.00268 -3.94
Average distance between inventors			0.00003 14.99
Generality			-0.00910 -1.50
Citation received in next 5 years after granted			0.00504 21.41
Log likelihood	-915540.04	-915538.44	-154134.47
Observations	566,306	566,306	88,996

Notes:

1. T-statistics in parentheses.
2. All models include fixed effects at the firm level.
3. All models include as regressors dummies for patent technology categories and for calendar years.
4. We include inventors whose address is in the US.
5. Inventors from patents applied for between 1975 and 2004, inclusive, only.

Table 5A
Determinants of Team Size

Dependent Var.: team size — Difference-in-Differences model

Variable	Model 4	Model 5	Model 6
Michigan*After_1986	0.11319 9.06	0.13241 5.39	0.10022 6.54
Michigan	-0.06774 -4.43	-0.07101 -3.20	-0.05145 -3.08
After_1986	0.32147 13.27	-0.03944 -0.77	0.18843 5.50
Log number of inventors within 50 miles	0.00689 3.64	0.00931 1.77	-0.00061 -0.25
Log R&D ($2009)	-0.03305 -7.37	0.05109 2.83	-0.02507 -4.69
Zero R&D dummy	0.00576 0.03	0.11127 0.13	0.05964 0.30
Log industry productivity ($2009)	-0.01714 -1.55	-0.10501 -1.58	-0.03491 -2.22
Log median wage for scientists ($2009)	0.07699 3.15	-0.06778 -1.05	0.13158 4.21
Log likelihood	-238878.22	-51927.54	-152166.64
Observations	145,970	32,752	94,479

Notes:

1. The specification of all models is Poisson model.
2. All models include fixed effects at the firm level.
3. All models include as regressors dummies for patent technology categories and for calendar years.
4. In model 5, we only include patents in Motor Vehicles industry.
5. In model 6, we employ the CEM method.
6. We include inventors whose address is in the US.
7. Inventors from patents applied for between 1975 and 2004, inclusive, only.
8. Control group includes only patents by firms in California.

Table 6
Determinants of Persistence of Pairs
Dep. Var.: number of repeats within 3 years of first appearance[1] Poisson model

Variable	Model 1	Model 2	Model 3
Log team size	-0.22982 -52.29	-0.23280 -52.96	-0.20647 -23.51
Noncompete	0.15685 19.10	0.59575 42.02	0.09422 5.52
T*Noncompete		-0.02549 -38.59	
Log number of inventors within 50 miles	0.09597 52.81	0.09149 50.43	0.05878 14.43
Log R&D ($2009)	-0.24412 -66.90	-0.25763 -70.10	-0.27028 -34.86
Zero R&D dummy	-1.63566 -5.41	-1.71063 -5.64	-0.27267 -0.38
Log industry productivity ($2009)	0.05321 5.22	0.05972 5.81	0.05045 2.42
Log median wage for scientists ($2009)	-0.12471 -5.92	-0.12946 -6.14	0.06020 1.22
Log number of claims			0.12985 27.65
Average patents per inventor in last 3 years			0.01430 14.96
Average distance between inventors			-0.00020 -31.65
Generality			-0.13524 -13.04
Citation received in next 5 years after granted			0.00312 30.91
Log likelihood	-781752.44	-780989.44	-175809.84
Observations	343,983	343,983	96,830

Notes:

1. First appearance during the period 1975 – 2001
2. T-statistics in parentheses.
3. All models include fixed effects at the firm level.
4. All models include as regressors dummies for patent technology categories and for calendar years.
5. We include inventors whose address is in the US.
6. Inventors from patents applied for between 1975 and 2004, inclusive, only.

Table 6A
Determinants of Persistence of Pairs
Dep. Var.: number of repeats within 3 years of first appearance[1] Diff-in-Diff model

Variable	Model 4	Model 5	Model 6
Log team size	-0.07141 -9.75	-0.33730 -15.48	-0.04825 -6.42
Michigan*After_1986	0.00486 0.16	0.25089 3.18	0.16784 4.58
Michigan	0.37503 9.67	-0.34811 -4.76	0.16068 3.76
After_1986	2.20027 43.16	0.16981 1.13	1.98547 32.05
Log number of inventors within 50 miles	0.04035 13.43	0.12900 10.58	0.03355 10.94
Log R&D ($2009)	-0.37766 -52.30	-1.25366 -35.62	-0.41320 -55.04
Zero R&D dummy	-2.10412 -2.36	-17.34469 -0.04	-2.30859 -2.57
Log industry productivity ($2009)	-0.35581 -16.16	1.21939 11.50	-0.37541 -15.25
Log median wage for scientists ($2009)	0.30605 7.48	-0.12616 -0.84	0.56520 12.98
Log likelihood	-241152.83	-38014.61	-221351.97
Observations	89,833	19,829	80,890

Notes:

1. First appearance during the period 1975 – 2001
2. The specification of all models is Poisson model.
3. All models include fixed effects at the firm level.
4. Models 5 and 6 include as regressors dummies for patent technology categories and for calendar years.
5. In model 5, we only include patents in Motor Vehicles industry.
6. In model 6, we employ the CEM method.
7. We include inventors whose address is in the US.
8. Inventors from patents applied for between 1975 and 2004, inclusive, only.

Appendix A: Dynamic Model of R&D Team Formation

A.1 Model Set-Up

We start with a firm which wishes to develop P (a strictly positive integer) ideas (or projects) into marketable products. The firm seeks to hire scientists for each project who are the only additional input in the development process. The firm operates for three periods. In the first period, scientists work on the project and develop a viable prototype. In the second period, the prototype is put in production and the firm earns revenue, without the aid of the scientists. By the end of the first period, each scientist possesses knowledge that enables him, if he desires, to leave for a rival firm and cash in on it. If the knowledge is leaked, the firm suffers a loss in revenue. At the beginning of the second period the firm and each scientist learn about the value of the scientist's knowledge to a rival. We assume that this 'external' value is a random variable, ρ_2 ($\in R$), and the density for ρ_2 is known a priori to the firm and the scientist. ρ_2 is the external value of the knowledge net of moving costs, which include the search cost of finding a suitable rival firm and any relocation expenses. For simplicity, all the scientists are intrinsically identical except for their research areas of specialization and are assumed to draw the same value of ρ_2 at the beginning of the second period. Therefore, if a single draw of ρ_2 turns out high enough, all scientists on each project leave. Otherwise, all stay. The firm launches a new set of P projects in the second period with those scientists from the first period if they decide to stay, or with a new group of scientists otherwise. In the third period the firm draws revenue from the new projects performed in the second period. In the third period, because no new products are developed all scientists leave at the period's beginning.

Following Becker and Murphy (1992), we assume that a continuum of tasks along an interval with length R must be performed for each project to develop a product. All P projects are symmetric in all aspects such as the task interval length and the revenue and cost functions (defined later) except that each project requires a different set of tasks. Moreover, we assume that the task interval for each project is a non-overlapping arc of a circle with the circumference of R·P. The circle is divided into P number of arcs each of which represents the required tasks for a project and the number of tasks required is the length R of the arc (see Figure A.1A for the example with P = 4).

With perfect complementarity among all tasks in the interval for a project, the revenue (Y) from a project takes the form of Leontief function: $Y = \min_{0 \leq t \leq R} Y(t)$. The rate of revenue from task t (Y(t)) equals the product of the working time devoted to task t (u(t)) and the output per unit of time spend on activity t (A(t)). All scientists are homogenous and each is specialized in a predetermined set of tasks. For example, a scientist is specialized in a broader set of tasks corresponding to biology while another is specialized in a narrower set of tasks corresponding to molecular biology. Each scientist allocates his total working time u uniformly on the tasks of his specialty. If the firm decides to hire n scientists to cover all the tasks required for a project, the firm seeks scientists in the labor market, each of whom should be specialized in the tasks with the interval length of (R/n), and therefore, $u(t) = u(R/n)^{-1}$. The productivity A(t) is assumed to depend positively on the degree of specialization: $A(t) = A(R/n)^{-\gamma}$, where A is a productivity parameter and γ (>0) denotes a specialization-intensity parameter. Note that the productivity is higher when each scientist is more specialized and the task interval he covers (R/n) is narrower. The revenue from a project is thus $Y = [u(R/n)^{-1}]\cdot[A(R/n)^{-\gamma}]$.

We recognize that the cost of coordinating the tasks among scientists in a project rises with the number of scientists. In particular, the coordination cost is assumed to take the form: $\Phi\cdot n^{\beta}$ where Φ is a cost parameter and β is a cost-intensity parameter. We assume β to be greater than 1, or increasing marginal cost of coordination with respect to the number of scientists, since the profit-maximizing solution may not exist otherwise.

The firm's loss of revenue due to knowledge leak is assumed to be a function of the number of scientists leaving (m) and the breadth (b) and depth (d) of each scientist's knowledge, $\Lambda(m, b, d)$. Revenue loss increases with the number of scientists leaving (i.e. $\partial\Lambda/\partial m>0$) and with the breadth or depth of knowledge (i.e. $\partial\Lambda/\partial b>0$, $\partial\Lambda/\partial d>0$). The number of scientists leaving has a scale effect on loss because more rival firms are likely to emerge. The effect of the breath on loss is because specialized scientists due to compartmentalization of research can transfer only fragmented knowledge of the project to a rival.

At the beginning of the second period, two types of uncertainty are resolved. The external value of knowledge transfer by a moving scientist (ρ_2) is revealed. If the scientist finds the external value sufficiently attractive, he leaves the firm. We assume that the random variable ρ_2 is equal to $\bar{\rho}_2 + e(b, d) + \varepsilon_2$ where ε_2 ($\in R$) is a mean zero random variable with density f, and $\bar{\rho}_2$ is a constant term. In this specification the external value is higher when the scientist leaves with wider breadth or more depth of knowledge, or $\partial e/\partial b>0$, $\partial e/\partial d>0$.

The second type of uncertainty that is resolved in the second period is the degree of technological change. Advances in technology may mean that the projects the firm pursued in the first period are no longer profit-maximizing in the second period. Our model assumes that the number of projects and the task circle given in the first period remain unchanged, and that the firm chooses the profit-maximizing project among projects with different sets of tasks whose revenue levels are revealed at the beginning of the second period. More specifically, we assume that the productivity parameter A can vary with different sets of tasks for projects on the circle. In Figure A.1B, we show the center location of tasks required for a project in the second period (denoted by r) on the horizontal axis and the corresponding level of parameter A on the vertical axis. In other words, a project with r requires all tasks within length R/2 on the left side and on the right side of r. The center location for the first-period project is by construction R/2. Our assumption is that there is a unique set of tasks that generates the maximum revenue which has the center location at r* and parameter A for alternative project locations is monotonically decreasing with the distance from the project location with the maximum A. Note that the level of A is defined only over the range of r in [0,R] since all projects are symmetric. We assume that r* ($\in$[0,R]) is a random variable and the density for r* is known a priori to the firm and the scientists. For analytical simplicity, we assume that the productivity parameter takes the form: $A = a_2 - a_1(r-r^*)^2$, where a_1 and a_2 are positive constants.

When the scientists stay in the second period, the only choice variable for the firm is the location of projects (that is, r) to maximize the profit. When the scientists stay, it is not generally optimal to choose r at r* (revenue-maximizing value) because the model recognizes the accumulation of team-specific human capital: coordination cost can be lowered if more from the same team of scientists in the first-period project are put together in the second period because those who worked together in the previous period have accumulated team-specific human capital. The coordination cost parameter Φ is shown in Figure A.1C to have the lowest point when r is R/2 (that is, when using the same group of scientists from the first-period project) and rise monotonically with the distance of r from the original set of tasks in the first period. More specifically, parameter Φ is assumed to take the functional form: $\Phi = \phi_2 - \phi_1 r(R-r)$ where ϕ_1 and ϕ_2 are positive constants. When the scientists from the first period stay, the firm chooses the projects in the second period by maximizing

(A.1) $[u(R/n)^{-1}]\cdot[A(R/n)^{-\gamma}].- \Phi\cdot n^{\beta}$,

with respect to r, where A and Φ are as defined above and n is predetermined from the expected profit maximization problem in the first period.

When the scientists leave the firm with better outside opportunity in the second period (or ρ_2 is sufficiently high), the firm hires a new group of scientists to maximize its profit in the remaining periods.[17]

The firm's objective is to maximize the expected profit from a single project since all P projects are intrinsically symmetric. The only choice variable for the firm in the first period is the number of scientists to hire (n). The expected profit in the first period is

[17] The second order condition of this maximization problem requires that $\beta>(1+\gamma)$.

(A.2) $-\ \phi_2 \cdot n^{\beta} - w_1 \cdot n + \delta \cdot u(R/n)^{-1} a_2 (R/n)^{-\gamma}$

$+\ \delta\ E\{-\Lambda(n,\ R/n,\ u(R/n)^{-1}) + \max_{s} [-\ \phi_2 \cdot s^{\beta} - w_{2a} \cdot s + \delta \cdot u(R/s)^{-1} a_2 (R/s)^{-\gamma} - \delta \cdot \Lambda(s,\ R/s,$

$u(R/s)^{-1})] \mid \rho_{2(\text{leaving})}, r^*\}$

$+\ \delta\ E\{ \max_{r} [-\ \Phi \cdot n^{\beta} - w_{2b} \cdot n + \delta \cdot u(R/n)^{-1} A (R/n)^{-\gamma} - \delta \cdot \Lambda(n,\ R/n,\ u(R/n)^{-1})] \mid \rho_{2(\text{staying})}, r^*\},$

where δ is a time discount factor and E is an expectation operator with respect to two random variables, ρ_2 and r*. The expectation operator in the second row is defined over the set of ρ_2 values such that the scientists leave, and the operator in the third row is defined over the set of ρ_2 values such that the scientists stay. w_1 is the wage in the first period, w_{2a} is the wage for those scientists newly hired in the second period, and w_{2b} is the wage in the second period for staying scientists. We will later show how these wages are endogenously determined in our model with the assumption that scientists receive the same wage rate regardless of the extent of specialization. A few points are noteworthy in this expression. First, the coordination costs in the first period and in the second period when scientists leave are associated with ϕ_2, the maximum level of cost parameter Φ, since team-specific human capital is never accumulated or it is lost due to mobility, respectively. Second, the productivity parameter A takes the maximum value a_2 in the first period and in the second period when scientists leave since no team-specific human capital is accumulated so that the firm only has to choose the revenue-maximizing projects in these two cases. Third, the number of new scientists hired in the second period (s) can be different from that of those hired in the first period (n) because the decision in the first period takes into account the chance of scientists staying.[18] Fourth, all scientists leave after the first period if the external value ρ_2, common for all scientists, is high enough. In the loss function $\Lambda(m, b, d)$, we thus have the number of leaving scientists (m) equal to n and the breadth (b) equal to R/n. The depth of knowledge (d) is assumed to be proportional to the working time devoted to each task, which is equal to $u(R/n)^{-1}$. Note that this loss function illustrates three distinct effects of the number of scientists in a team on revenue loss. First, more scientists on the team implies more rival firms, should the scientists leave. Second, more scientists results in greater specialization, and hence more knowledge depth, which makes for more effective rivals. Third, more scientists implies each scientist commands a smaller slice of the big picture, thus limiting their effectiveness as a rival.

We assume that the scientist like the firm is risk neutral and therefore maximizes his expected income. The scientist chooses at the beginning of the first period whether to accept the firm's offer or to work for another firm. To simplify the analysis we assume that outside the firm he would acquire no appropriable proprietary knowledge but would receive his marginal product or reservation wage, w_r, in the first, the second, or the third period. The firm's offer consists of a guaranteed first period wage, w_1, and a second period wage, w_{2b}, when the scientist stays in the second period. The firm specifies the second period wage only after the random variables ρ_2 and r* are realized, taking the scientist's decision in the second period as given. If the scientist accepts the job offer in the first period, at the beginning of the second period he chooses among two options based on the realized ρ_2 and r*. He may join a rival, performing work equal in value to w_r, and, in addition, selling his knowledge and receiving its full

[18] If scientists always leave after the first period, n should be equal to s.

value, ρ_2. His third-period earning is w_r in this case. (Recall that everyone leaves the firm at the end of the second period.) He may remain with the firm in the second period, earning w_{2b} and performing work equal in value to w_r. He earns $E(\rho_3|b, d)$ plus w_r in the third period in this case where $E(\rho_3|b, d)$ is the expected external gain from knowledge he acquires during the second period which is a function of knowledge breadth and depth, b and d.

In the event the scientist stays in the second period because the draw of ρ_2 is lower than the threshold value (see equation A.6), he will be paid at least as much as he would get outside. This implies that $w_{2b} + \delta\, E(\rho_3|b, d) + \delta\, w_r = \rho_2 + w_r + \delta\, w_r$, where the left-hand side is the expected earning if the scientist stays and the right-hand side is the expected earnings if he moves. Therefore, wage w_{2b} is given by

(A.3) $w_{2b} = \rho_2 + w_r - \delta\, E(\rho_3|R/n, u(R/n)^{-1})$.

Note that the breadth of knowledge is R/n and the depth is $u(R/n)^{-1}$ since there are n scientists in the team. The scientist accepts the contract in the first period if the expected earnings in three periods are at least as much as $(1+\delta+\delta^2)w_r$:

$$(1+\delta+\delta^2)w_r = w_1 + \delta E\{\rho_2+w_r+\delta w_r|\rho_{2(\text{leaving})}, r^*\}$$

$$+ \delta E\{w_{2b}+\delta E(\rho_3|R/n, u(R/n)^{-1})+\delta w_r|\rho_{2(\text{staying})}, r^*\}, \text{ or}$$

(4) $w_1 = w_r - \delta\, \bar{\rho}_2 - \delta\, e(R/n, u(R/n)^{-1})$.

Note that the breadth and the depth of knowledge in function e are again R/n and $u(R/n)^{-1}$, respectively.

When the firm hires new scientists in the second period, the wage it pays should be such that the expected earnings for a scientist from accepting an offer are equal to the expected earnings otherwise: $w_{2a} + \delta\, E(\rho_3|R/s,\, u(R/s)^{-1}) + \delta\, w_r = w_r + \delta\, w_r$. This implies that

$$\text{(A.5)} \quad w_{2a} = w_r - \delta\, E(\rho_3|R/s,\, u(R/s)^{-1}).$$

In equations (A.3) and (A.5), w_{2a} is different from w_{2b} in two ways. First, w_{2b} includes the realized value of ρ_2 because the firm needs to match the wage to what the scientist can earn outside. Second, the expected external values from knowledge transfer in the third period and thus the two wages, w_{2a} and w_{2b}, can differ because n is not equal to s in general.

A.2 Job Turnover among Scientists

For a scientist, the extra return from leaving the firm instead of staying is $[\rho_2 + w_r + \delta\, w_r] - [w_{2b} + \delta\, E(\rho_3|R/n,\, u(R/n)^{-1}) + \delta\, w_r]$. If this extra return exceeds the maximum the firm is willing to pay per scientist to keep the scientists, which is the difference in the firm's profit between when the scientists stay and when they leave, the firm will let the scientists go. The condition for mobility is therefore,

$$\rho_2 + w_r - w_{2b} - \delta\, E(\rho_3|R/n,\, u(R/n)^{-1})$$

$$> (1/n) \max_r\, [-\,\Phi\cdot n^\beta - w_{2b}\cdot n + \delta\cdot u(R/n)^{-1} A(R/n)^{-\gamma} - \delta\cdot\Lambda(n,\, R/n,\, u(R/n)^{-1})]$$

$$-\,(1/n) \max_s\, [-\,\phi_2\cdot s^\beta - w_{2a}\cdot s + \delta\cdot u(R/s)^{-1} a_2(R/s)^{-\gamma} - \delta\cdot\Lambda(s,\, R/s,\, u(R/s)^{-1})]$$

$$+\,(1/n)\Lambda(n,\, R/n,\, u(R/n)^{-1}), \text{ or}$$

$$\text{(A.6)} \quad \varepsilon_2 > \delta\, E(\rho_3|R/n,\, u(R/n)^{-1}) - w_r - e(R/n,\, u(R/n)^{-1}) - \bar{\rho}_2 + (1/n)\Lambda(n,\, R/n,\, u(R/n)^{-1})$$

$$+\,(1/n) \max_r\, [-\,\Phi\cdot n^\beta + \delta\cdot u(R/n)^{-1} A(R/n)^{-\gamma} - \delta\cdot\Lambda(n,\, R/n,\, u(R/n)^{-1})]$$

$$-\ (1/n)\max_{s}\ [-\ \phi_2 \cdot s^{\beta} - w_{2a} \cdot s + \delta \cdot u(R/s)^{-1} a_2 (R/s)^{-\gamma} - \delta \cdot \Lambda(s, R/s, u(R/s)^{-1})]\ \equiv\ \bar{\varepsilon}_2\,.$$

Intuitively, the scientists are more likely to leave if the external opportunity is more attractive on average ($\bar{\rho}_2$ or e is higher), the loss to the firm in the second period is smaller (Λ in the first row is smaller), the firm's gain due to team-specific human capital accumulation is smaller (ϕ_1 is lower), or the technology in the second period is changed more so that the task set is more different from that in the first period (r* deviates on average farther from R/2). The reason for the last effect is that the farther r* is from R/2 the lower the firm's profit when the scientists stay and this will deter the firm from keeping the scientists. As the difference between the revenue from adopting a new revenue-maximizing project and from staying with the old project rises (larger a_1), the scientists are more likely to leave since the firm has a smaller incentive to utilize the accumulated team-specific human capital.[19]

The effects of ϕ_2, a_2, and w_r depend on whether n is bigger than the optimal s. For instance, an increase in the reservation wage will raise both wages w_{2a} and w_{2b} in the second period by the same amount. If the optimal n is smaller than the optimal s (which is consistent with the observed increase in team size), the total wage bill in the case when the scientists stay will rise less, which makes mobility less likely. If the optimal team size grows over time, lower expropriation loss or higher productivity provides firms a stronger incentive to hire a new group of scientists and thus raises mobility.

The effect of the project range R on mobility is ambiguous. An increase in

[19] Derivations of the comparative statics results in this and the following sections are reported in A.6.

project range R will increase the breadth of knowledge each scientist can acquire. This will in turn increase the external value for scientists and make them more likely to leave. On the other hand, the expropriation loss will rise and the firm will want to retain scientists.

A.3 Choice of Projects

Persistent teamwork among scientists can be determined by two factors in our model. Teamwork will be less persistent if more scientists leave, provided that the cooperation among the scientists leaving is less likely when they leave the firm. Secondly, if the firm chooses the second-period projects with similar task sets to those in the first period, the continuation of teamwork among the scientists who stay will be persistent. We investigate the determinants for the second factor in this section. The net effect of these two factors will be discussed later in section A.5.

From the first-order optimality condition for the problem in equation (1), we can derive the optimal value of project location, r:

(7) $r = \Omega\, r^* + (1-\Omega)\,(R/2)$,

where $\Omega \equiv u \cdot a_1 (R/n)^{-1-\gamma} / [u \cdot a_1 (R/n)^{-1-\gamma} + \phi_1 \cdot n^{\beta}]$ (<1).

Equation (7) implies that the optimal r is located between the revenue-maximizing value of r (r*) and the coordination-cost-minimizing value of r (R/2). If the tasks required for new technology in the second period are more similar to those in the first period, the optimal r should be closer to R/2 and more scientists are teamed up again. In other words, teamwork is less persistent when r* deviates farther from R/2. Other comparative statics analyses include:

(a) If the marginal effect of team-specific human capital on reducing coordination cost is more pronounced (higher ϕ_1), Ω gets smaller, which implies that more scientists are teamed up again.

(b) If the increase in revenue by adopting a new revenue-maximizing project instead of staying with the old project is less (smaller a_1), Ω gets smaller, and more scientists are teamed up again.

(c) The project range, R, has two opposing effects on the optimal r. Given the number of scientists, wider range lowers Ω, which implies that the optimal r is closer to R/2. On the other hand, wider range makes wider the distance in tasks between two scientists, which makes the optimal r farther from R/2.

A.4 Team size

Taking wages given, the firm chooses team size (n) to maximize profit. The first-order optimization condition with respect to team size (n) is

$$(8) \qquad \phi_2 \cdot \beta \cdot n^{\beta-1} + w_1 \;=\; \delta \cdot u \cdot a_2 (R)^{-1-\gamma} (1+\gamma) n^{\gamma} + \delta \frac{\partial}{\partial n} E\{-\Lambda(n, R/n, u(R/n)^{-1}) \mid \rho_{2(leaving)}, r^*\}$$

$$+\; \delta \frac{\partial}{\partial n} E\{ \max_{r} [-\; \Phi \cdot n^{\beta} - w_{2b} \cdot n + \delta \cdot u (R/n)^{-1} A (R/n)^{-\gamma} - \delta \cdot \Lambda(n, R/n, u(R/n)^{-1})] \mid \rho_{2(staying)}, r^*\}.$$

The left hand side is the marginal cost of n in coordination and wage cost, respectively. The right hand side shows the marginal benefits of n. The first term on this side is the marginal benefit due to specialization. The second term reflects how the loss can be minimized by choosing the right team size. The last term is related to the expected marginal profit of n from the second period project when the scientists stay, which is not necessarily zero since the optimal n is maximizing not just the profit from the second period project. Some comparative statics analyses are in order.

(a) An increase in the external value $\bar{\rho}_2$ can have a positive effect on n through a couple of channels. Corresponding to the effect of general training in labor economics, better outside opportunity lowers the wage in the first period, and thus the marginal cost of n, which increases the optimal team size. (We can show that this effect always overwhelms the opposing effect of rising wage w_{2b} due to an increase in $\bar{\rho}_2$). In addition, when mobility is more likely with higher $\bar{\rho}_2$, the marginal benefit of raising n to reduce the breadth of knowledge each scientist acquires (or, to compartmentalize research operation) becomes more significant. This will also lead to bigger team size. However, rising $\bar{\rho}_2$ may affect n adversely through other channels. With the scale effect of n on expropriation loss, rising number of departing scientists with increasing $\bar{\rho}_2$ will tend to raise cost and reduce n. The firm also has an incentive to reduce n and hire scientists with less depth in knowledge to lower the appropriation loss with mobility. Additionally, suppose the optimal n is smaller than the optimal number of scientists (s) that maximizes the profit only in the second-period project when the scientists stay so that the marginal profit of n is strictly positive in terms of the profit from the second-period project.[20] This can take place if the marginal cost of n in the first period project is sufficiently high. If the probability that the scientists stay falls due to rising $\bar{\rho}_2$, the expected marginal profit from the second-period project when

[20] Note that the optimal s is known in period 1 since no random variable is involved in the maximization problem for s (see equation 1).

scientists stay falls and thus the optimal n falls. The last effect evidently will be opposite or none if the optimal n is bigger than, or equal to the optimal s. (See Appendix for details.)

(b) The direct effect of higher marginal productivity in research (a_2) or lower marginal cost in coordination (ϕ_2) will be to increase team size. A change in a_2 or ϕ_2 can also affect the optimal team size by altering the probability of mobility. However, the effect of a_2 or ϕ_2 on the probability of mobility is ambiguous as seen in section 2.1. In our empirical study, we expect to only observe the direct effect of a_2 or ϕ_2 if mobility is controlled for in our specification.

(c) Higher ϕ_1 implies that more team-specific human capital can be accumulated through teamwork, and coordination cost thus can be reduced further when the scientists stay. On one hand, lower coordination costs can lead to an increase in team size. On the other hand, reduced coordination costs when the scientists stay raises the firm's incentive to retain the scientists. This will reduce labor mobility and the firm's need to compartmentalize research operation, which will lead to smaller team size. If we can control for scientists' mobility in the empirical analysis, parameter ϕ_1 is expected to have a positive effect on n.

(d) Wages in the first and second periods rise with the reservation wage w_r. An increase in w_r will therefore have an adverse effect on team size. A change in w_r, however, can also affect the optimal team size through altering the probability of mobility that can be affected ambiguously by the reservation wage w_r. With mobility controlled for in the empirical specification, the reservation wage is thus expected to have a negative effect on n.

(e) When the technology in the second period is more drastically changed from the first period so that r* deviates farther from R/2, the marginal productivity (A) falls, which will tend to decrease the team size. Section A.2 shows that scientists are more likely to move in this case. The firm will then increase or decrease the team size. When mobility is controlled for in the empirical specification, we expect an adverse effect of technological change on n.

(f) Setting aside the ambiguous effect of parameter R on the probability of mobility and thus the team size, an increase in R can be shown to have at least two opposing effects. Since the breadth of knowledge each scientist can acquire will increase with R, the scientists are expected to have more to appropriate when leaving, which will lower wage w_1 and increase the team size. On the other hand, the firm will experience a reduction in the gains from specialization in the first and second periods. This will decrease the team size.

A.5 Persistent teamwork

Teamwork persistence is shaped via two channels in our model: scientist mobility (section A.2) and project overlap (section A.3). According to the former channel, teamwork will be less persistent if more scientists leave. As for the latter channel, teamwork will be more persistent if the firm chooses the second-period projects with task sets similar to those in the first period.

(a) The discussions in the previous sections suggest that labor mobility due to rising external value of transmitted knowledge ($\bar{\rho}_2$) can produce at least two opposing effects on the persistence of teamwork. Mobility directly reduces teamwork persistency since scientists are less likely to work together again when they depart the firm. On the other hand, if team size becomes bigger as labor mobility rises, possibly due to the firm's motivation to compartmentalize research operation with higher mobility, the firm will choose the projects in the second period that are similar to those in the first period in order to take advantage of accumulated team-specific human capital. This will raise teamwork persistency. However, with team size held constant, we will only have the former, negative effect. Our aim is to generate testable empirical implications. Because empirical researchers should be able to control for team size, in the following comparative statics analyses, we ignore the parameter effects on team persistency through team size since our empirical analysis will condition on team size.

(b) Parameters ϕ_2, a_2, and w_r can affect persistency only through the scientist mobility channel, and their effects depend on whether n is bigger than the optimal s or not. If the optimal n is smaller than the optimal s, an increase in a_2, or a decrease in ϕ_2, and w_r will make mobility more likely.

(c) Higher efficiency in team-specific human capital accumulation (ϕ_1) will induce the firm to choose the similar projects in the second period as in the first. This will increase teamwork persistency when scientists stay. In addition, rising ϕ_1 will result in reduced labor mobility, which also raises the persistence of teamwork.

(d) When technology is changed rapidly so that the average r* deviates farther from R/2, the firm's profit is reduced in case when scientists stay. This will increase mobility and reduce the persistency of teamwork. Furthermore, the optimal location of projects in the second period (r) will get far away from that in the first period, which will reduce the persistency of teamwork.

(e) The project range R has an ambiguous effect on scientist mobility (see section A.2) and on project overlap (see section A.3).

(f) An exogenous increase in the expropriation loss will increase the persistency of teamwork through the channel of scientist mobility.

A.6 Proofs of comparative static results

1. Comparative statics analysis in section "A.2 Job Turnover among Scientists"

In order to know how a parameter in the model influences mobility, we only have to show whether the parameter increases or decreases $\bar{\varepsilon}_2$ in equation 6. If it increases $\bar{\varepsilon}_2$, mobility is less likely and vice versa.

$\partial \bar{\varepsilon}_2 / \partial \bar{\rho}_2 = -1 < 0.$

$\partial \bar{\varepsilon}_2 / \partial \phi_2 = (-n^{\beta} + s^{\beta}) / n \quad > 0$ if $n < s$,
< 0 if $n > s$.

$\partial \bar{\varepsilon}_2 / \partial \phi_1 = -n^{\beta-1} r (r - R) > 0.$

$\partial \bar{\varepsilon}_2 / \partial a_1 = \delta u (R/n)^{-1-\gamma} [-(r - r^*)^2] / n < 0$

$\partial \bar{\varepsilon}_2 / \partial a_2 = u(R)^{-1-\gamma} (n^{1+\gamma} - s^{1+\gamma}) / n \quad > 0$ if $n > s$,
< 0 if $n < s$.

$\partial \bar{\varepsilon}_2 / \partial w_r = -1 + s/n \quad > 0$ if $n < s$,
< 0 if $n > s$.

$\partial \bar{\varepsilon}_2 / \partial |r^* - R/2| < 0$

$\partial \bar{\varepsilon}_2 / \partial R = -(\partial e / \partial b)/n + (\partial e / \partial d) [\, u(R/n)^{-2}/n] + \ldots \gtrless 0$

2. Comparative statics analysis in section "A.4 Team size"

Rearranging equation 8, we get

$$0 = -\phi_2 \cdot \beta \cdot n^{\beta-1} - w_1 + \delta \cdot u \cdot a_2 (R)^{-1-\gamma} (1+\gamma) n^{\gamma} + \delta \frac{\partial}{\partial n} E\{-\Lambda(n, R/n, u(R/n)^{-1}) \mid \rho_{2(\text{leaving})}, r^*\}$$
$$+ \delta \frac{\partial}{\partial n} E\{ \max_r [-\Phi \cdot n^{\beta} - w_{2b} \cdot n + \delta \cdot u (R/n)^{-1} A (R/n)^{-\gamma} - \delta \cdot \Lambda(n, R/n, u(R/n)^{-1})] \mid \rho_{2(\text{staying})}, r^*\}$$
$$\equiv \Pi .$$

Taking the total differentiation of this equation, we have

$$0 = (\partial\Pi/\partial n)dn + (\partial\Pi/\partial \bar{\rho}_2) d\bar{\rho}_2 + (\partial\Pi/\partial a_2)da_2 + (\partial\Pi/\partial\phi_2)d\phi_2 + (\partial\Pi/\partial\phi_1)d\phi_1 + (\partial\Pi/\partial w_r)dw_r$$
$$+ (\partial\Pi/\partial r^*)dr^* + (\partial\Pi/\partial R)dR,$$

where $\partial\Pi/\partial n < 0$ as long as the second-order optimality condition holds. The partial effect of $\bar{\rho}_2$, for example, will be then $dn/d\bar{\rho}_2 = -(\partial\Pi/\partial \bar{\rho}_2) / (\partial\Pi/\partial n)$.

(a) $\partial\Pi/\partial\bar{\rho}_2 = \left[-\partial w_1/\partial\bar{\rho}_2 - \delta\int_0^R\int_{-\infty}^{\bar{\varepsilon}_2} f(\varepsilon)g(r)d\varepsilon dr\right]$
$+\left[-\delta\int_0^R \frac{\partial}{\partial n}\Lambda(n, R/n, u(R/n)^{-1})\ f(\bar{\varepsilon}_2)g(r)dr\right]$
$+\quad\left[-\delta\int_0^R\left\{\frac{\partial}{\partial n}\left(-\Phi n^{\beta} + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) + \delta nE(\rho_3)\right) - e(.) - \bar{\rho}_2 - \bar{\varepsilon}_2 - w_r + \delta E(\rho_3)\right\}\ f(\bar{\varepsilon}_2)g(r)dr\right]$
$+\quad\left[-\delta\int_0^R\Lambda(.)\frac{\partial\bar{\varepsilon}_2}{\partial n}f(\bar{\varepsilon}_2)g(r)dr - \delta\int_0^R\left(-\Phi n^{\beta} + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) - n\bar{\rho}_2 - n\bar{\varepsilon}_2 - w_r n + \delta nE(\rho_3)\right)\frac{\partial\bar{\varepsilon}_2}{\partial n}\ f'(\bar{\varepsilon}_2)g(r)dr\right]$
= [A1] + [A2] + [A3] + [A4]
where $\partial w_1/\partial\bar{\rho}_2 = -1$.

Term A1 shows a direct effect of $\bar{\rho}_2$: better outside opportunity lowers the first-period wage. This wage effect (the first term in A1) always dominates rising second-period wage (the second term in A1). All other three terms A2-A4 pertain to indirect effects of $\bar{\rho}_2$ via labor mobility. Term A2 is the influence through expropriation loss when $\bar{\rho}_2$ affects the probability of mobility. This term can be negative if the scale plus the depth effect is stronger, or positive if the compartmentalization effect is more dominant. Term A3 refers to the last effect mentioned in (a) in section 3.3. Term A4 is related to how the marginal effect of n on the probability of moving is affected by $\bar{\rho}_2$, which is ambiguous. All the following derivations have the four channels similar to this analysis.

(b) $\partial\Pi/\partial a_2 = [u(R)^{-1-\gamma}(1+\gamma)n^{\gamma}\,(1+\delta\int_0^R\int_{-\infty}^{\bar{\varepsilon}_2} f(\varepsilon)g(r)d\varepsilon dr)]$
$+ [\delta\int_0^R\frac{\partial}{\partial n}\Lambda(.)\frac{\partial\bar{\varepsilon}_2}{\partial a_2}\ f(\bar{\varepsilon}_2)g(r)dr]$
$+ [\delta\int_0^R\left\{\frac{\partial}{\partial n}\left(-\Phi n^{\beta} + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) + \delta nE(\rho_3)\right) - e(.) - \bar{\rho}_2 - \bar{\varepsilon}_2 - w_r + \delta E(\rho_3)\right\}\frac{\partial\bar{\varepsilon}_2}{\partial a_2}f(\bar{\varepsilon}_2)g(r)dr]$
$+[\delta\int_0^R\Lambda(.)\left(\frac{\partial^2\bar{\varepsilon}_2}{\partial n\,\partial a_2}f(\bar{\varepsilon}_2) + \frac{\partial\bar{\varepsilon}_2}{\partial n}f'(\bar{\varepsilon}_2)\frac{\partial\bar{\varepsilon}_2}{\partial a_2}\right)g(r)dr + \delta\int_0^R\left(u(R/n)^{-1-\gamma} - n\frac{\partial\bar{\varepsilon}_2}{\partial a_2}\right)\frac{\partial\bar{\varepsilon}_2}{\partial n}f(\bar{\varepsilon}_2)g(r)dr + \delta\int_0^R\left(-\Phi n^{\beta} + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) - n\bar{\rho}_2 - n\bar{\varepsilon}_2 - w_r n + \delta nE(\rho_3)\right)\left(\frac{\partial^2\bar{\varepsilon}_2}{\partial n\,\partial a_2}f(\bar{\varepsilon}_2) + \frac{\partial\bar{\varepsilon}_2}{\partial n}f'(\bar{\varepsilon}_2)\frac{\partial\bar{\varepsilon}_2}{\partial a_2}\right)g(r)dr]$
= [B1] + [B2] + [B3] + [B4]
where $\frac{\partial\bar{\varepsilon}_2}{\partial a_2}$ is shown in Appendix 1, and $\frac{\partial^2\bar{\varepsilon}_2}{\partial n\,\partial a_2} = u(R/n)^{-1-\gamma}(\gamma n^{\gamma-1} + s^{1+\gamma}/n^2) > 0$.

Term B1 pertains to the improved productivity in the first and second period when a_2 rises. This term is negative. Term B2 is the influence through expropriation loss when an increase in a_2 affects the probability of mobility. Term B3 is about the effect of a_2 via mobility on the marginal profit of n in the second period. Term B4 is related with how the marginal effect of n on the probability of mobility is affected by a_2.

(c) $\partial\Pi/\partial\phi_2 = [-\beta n^{\beta-1}\,(1+\delta\int_0^R\int_{-\infty}^{\bar{\varepsilon}_2} f(\varepsilon)g(r)d\varepsilon dr)]$
$+ [\delta\int_0^R\frac{\partial}{\partial n}\Lambda(.)\frac{\partial\bar{\varepsilon}_2}{\partial\phi_2}\ f(\bar{\varepsilon}_2)g(r)dr]$

$+ \ [\delta \int_0^R \left\{\frac{\partial}{\partial n}\left(-\Phi n^\beta + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) + \delta n E(\rho_3)\right) - e(.) - \bar{\rho}_2 - \bar{\varepsilon}_2 - w_r + \delta E(\rho_3)\right\} \frac{\partial \bar{\varepsilon}_2}{\partial \phi_2} f(\bar{\varepsilon}_2) g(r) dr]$

$+ \ [\delta \int_0^R \Lambda(.) \left(\frac{\partial^2 \bar{\varepsilon}_2}{\partial n \, \partial \phi_2} f(\bar{\varepsilon}_2) + \frac{\partial \bar{\varepsilon}}{\partial n} f'(\bar{\varepsilon}_2) \frac{\partial \bar{\varepsilon}_2}{\partial \phi_2}\right) g(r) dr \ - \ \delta \int_0^R \left(n^\beta + n \ \frac{\partial \bar{\varepsilon}_2}{\partial \phi_2}\right) \frac{\partial \bar{\varepsilon}_2}{\partial n} f(\bar{\varepsilon}_2) g(r) dr$
$+ \delta \int_0^R \left(-\Phi n^\beta + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) - n\bar{\rho}_2 - n\bar{\varepsilon}_2 - w_r n + \delta n E(\rho_3)\right) \left(\frac{\partial^2 \bar{\varepsilon}_2}{\partial n \, \partial \phi_2} f(\bar{\varepsilon}_2) + \frac{\partial \bar{\varepsilon}_2}{\partial n} f'(\bar{\varepsilon}_2) \frac{\partial \bar{\varepsilon}_2}{\partial \phi_2}\right) g(r) dr]$

= [C1] + [C2] + [C3] + [C4]

Term C1 is negative because of rising coordination cost in the first and second period when ϕ_2 rises. Term C2 is the influence through expropriation loss when an increase in ϕ_2 affects the probability of mobility. Term C3 is about the effect of ϕ_2 via mobility on the marginal profit of n in the second period. Term C4 is related with how the marginal effect of n on the probability of mobility is affected by ϕ_2.

(d) $\partial\Pi/\partial\phi_1 = [\delta\int_0^R \int_{-\infty}^{\bar{\varepsilon}_2} \left(-\frac{\partial^2 \Phi}{\partial n \, \partial \phi_1} n^\beta - r(r-R)\beta n^{\beta-1}\right) f(\varepsilon) g(r) d\varepsilon dr)]$

$+ [\delta\int_0^R \frac{\partial}{\partial n} \Lambda(.) \frac{\partial \bar{\varepsilon}_2}{\partial \phi_1} \ f(\bar{\varepsilon}_2) g(r) dr]$

$+ \ [\delta \int_0^R \left\{\frac{\partial}{\partial n}\left(-\Phi n^\beta + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) + \delta n E(\rho_3)\right) - e(.) - \bar{\rho}_2 - \bar{\varepsilon}_2 - w_r + \delta E(\rho_3)\right\} \frac{\partial \bar{\varepsilon}_2}{\partial \phi_1} f(\bar{\varepsilon}_2) g(r) dr]$

$+ \ [\delta \int_0^R \Lambda(.) \left(\frac{\partial^2 \bar{\varepsilon}_2}{\partial n \, \partial \phi_1} f(\bar{\varepsilon}_2) + \frac{\partial \bar{\varepsilon}}{\partial n} f'(\bar{\varepsilon}_2) \frac{\partial \bar{\varepsilon}_2}{\partial \phi_1}\right) g(r) dr \ - \ \delta \int_0^R \left(r(r-R) n^\beta + n \ \frac{\partial \bar{\varepsilon}_2}{\partial \phi_1}\right) \frac{\partial \bar{\varepsilon}_2}{\partial n} f(\bar{\varepsilon}_2) g(r) dr \ + \delta \int_0^R \left(-\Phi n^\beta + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) - n\bar{\rho}_2 - n\bar{\varepsilon}_2 - w_r n + \delta n E(\rho_3)\right) \left(\frac{\partial^2 \bar{\varepsilon}_2}{\partial n \, \partial \phi_1} f(\bar{\varepsilon}_2) + \frac{\partial \bar{\varepsilon}_2}{\partial n} f'(\bar{\varepsilon}_2) \frac{\partial \bar{\varepsilon}_2}{\partial \phi_1}\right) g(r) dr]$

= [D1] + [D2] + [D3] + [D4]

where $\frac{\partial^2 \Phi}{\partial n \, \partial \phi_1} = 2\Omega(r^* - R/2)^2 \frac{\partial \Omega}{\partial n} < 0$.

Term D1 is positive because an increase in ϕ_1 raises the benefit of team-specific human capital accumulation and thus increases team size. Terms D2-D4 are related to the effect of ϕ_1 via labor mobility.

(e) $\partial\Pi/\partial w_r = [-\partial w_1/\partial w_r - \delta\int_0^R \int_{-\infty}^{\bar{\varepsilon}_2} f(\varepsilon) g(r) d\varepsilon dr]$

$+ [\delta\int_0^R \frac{\partial}{\partial n} \Lambda(.) \frac{\partial \bar{\varepsilon}_2}{\partial w_r} \ f(\bar{\varepsilon}_2) g(r) dr]$

$+ \ [\delta \int_0^R \left\{\frac{\partial}{\partial n}\left(-\Phi n^\beta + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) + \delta n E(\rho_3)\right) - e(.) - \bar{\rho}_2 - \bar{\varepsilon}_2 - w_r + \delta E(\rho_3)\right\} \frac{\partial \bar{\varepsilon}_2}{\partial w_r} f(\bar{\varepsilon}_2) g(r) dr]$

$+ \ [\delta \int_0^R \Lambda(.) \left(\frac{\partial^2 \bar{\varepsilon}_2}{\partial n \, \partial w_r} f(\bar{\varepsilon}_2) + \frac{\partial \bar{\varepsilon}_2}{\partial n} f'(\bar{\varepsilon}_2) \frac{\partial \bar{\varepsilon}_2}{\partial w_r}\right) g(r) dr \ - \ \delta \int_0^R \left(1 + n \ \frac{\partial \bar{\varepsilon}_2}{\partial w_r}\right) \frac{\partial \bar{\varepsilon}_2}{\partial n} f(\bar{\varepsilon}_2) g(r) dr$
$+ \delta \int_0^R \left(-\Phi n^\beta + uA(R/n)^{-1-\gamma} - \delta\Lambda(.) - ne(.) - n\bar{\rho}_2 - n\bar{\varepsilon}_2 - w_r n + \delta n E(\rho_3)\right) \left(\frac{\partial^2 \bar{\varepsilon}_2}{\partial n \, \partial w_r} f(\bar{\varepsilon}_2) + \frac{\partial \bar{\varepsilon}_2}{\partial n} f'(\bar{\varepsilon}_2) \frac{\partial \bar{\varepsilon}_2}{\partial w_r}\right) g(r) dr]$

= [E1] + [E2] + [E3] + [E4]

where $\partial w_1/\partial w_r = 1$.
Term E1 is negative because an increase in the reservation wage raises wages in both periods and thus reduces team size. Terms D2-D4 are related to the effect of w_r via labor mobility.

(f) $\partial\Pi/\partial R = [-\ \partial w_1/\partial R] + [-\ ua_2(R)^{-2-\gamma}(1+\gamma)^2 n^{\gamma}\ ] + ...$
$= [F1] + [F2] + \ldots$
We have more channels besides terms F1 and F2 through which R can affect n. Term F1 can be either positive or negative while F2 is negative. Term F1 is positive if the compartmentalization effect is dominant. Term F2 is negative since rising R lowers the benefit of specialization.

Figure A.1A

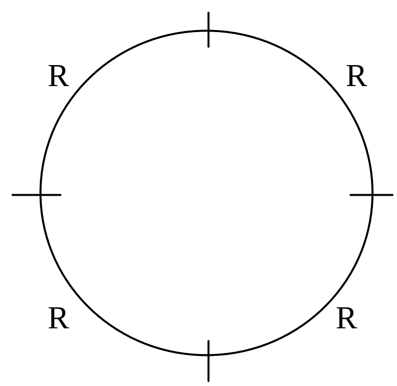

Figure A.1B

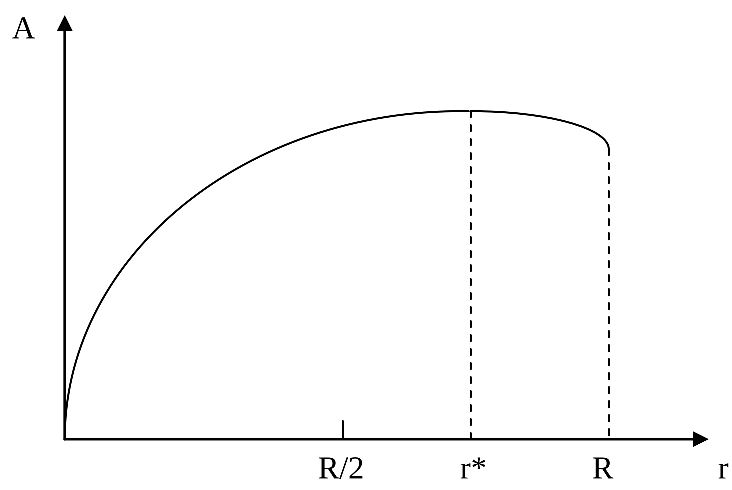

Figure A.1C

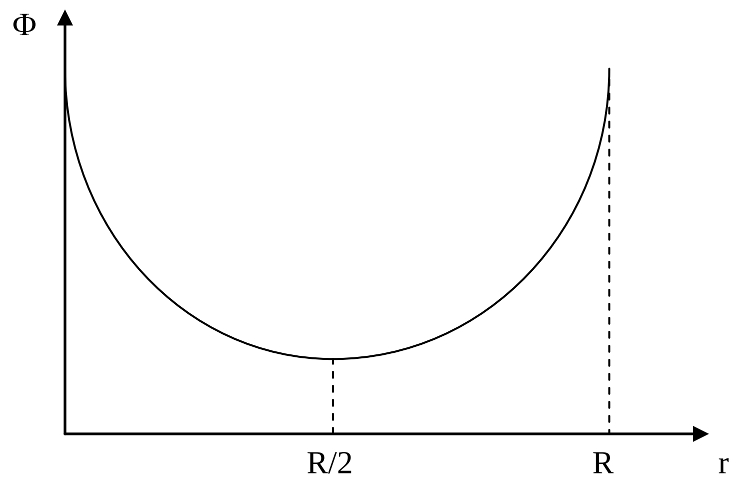

Figure A.1D