May 2019

ESSAYS ON URBAN ECONOMICS, SOCIAL NETWORKS AND INNOVATION

Yanmin Yang
Syracuse University

ABSTRACT

My dissertation comprises two papers on urban economics, social networks, and innovation. The main question that connects these two papers is what the trade-off between inventor teams and social networks in patent development is. Two papers provide empirical evidence and answer the question from different perspectives.

Chapter 1 examines the joint effect of inventor social networks and team size on patent impact. Using data from the United States Patent and Trademark Office (USPTO) from 1975 to 2010, results confirm that the marginal effect of network size on patent impact decreases as team size expands, indicating that inventor networks and team size are substitutes. The substitution effect is stronger for teams and networks within a shared technology field. The marginal effect of inventor networks on the amount of knowledge and the speed with which new knowledge in a field is adopted also diminishes as team size increases.

Chapter 2 identifies the causal effect of non-compete clause enforceability on patent inventor team size, external network size and accessibility using a quasi-natural experiment of the Michigan Antitrust Reform Act (MARA) in 1985. Compared with states with similar enforceability from 1980 to 1990, patents in Michigan are 4% less likely developed by single inventors and 1.3% more likely produced by large teams with at least four inventors. Old firms, especially those relying heavily on hiring inventors from other firms for patent development, experience a fast expansion in inventor team size in the short run. Meanwhile, patent development in Michigan is 2.6% less likely to access external knowledge and information through inventors recently hired from other firms within the same state. However, there is no clear relationship between MARA and the external network sizes. Most of these findings are in support of the results of Chapter 1.

ESSAYS ON URBAN ECONOMICS, SOCIAL NETWORKS AND INNOVATION

by

Yanmin Yang

B.A., Fudan University, 2011
M.A., Fudan University, 2014

Dissertation

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in *Economics*.

Syracuse University

May 2019

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Professor Stuart Rosenthal for his guidance of my PhD study. I appreciate the insightful comments from Professors William Horrace, Kristy Buzard, Alfonso Flores-Lagunes and Alexander Rothenberg. I am also grateful for suggestions received from participants at the American Real Estate and Urban Economics Association (AREUEA) poster session at the 2019 ASSA. Lastly, I want to thank my mother Min Chen, my father Yuezhong Yang, as well as my grandparents, for their support of my PhD study abroad.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Innovation in a Networked World: Inventor Teams, Social Networks and Patent Impact

Chapter 2: Non-Compete Clauses and R&D Inventor Team Size

LIST OF TABLES

Chapter 1: Innovation in a Networked World: Inventor Teams, Social Networks and Patent Impact

Chapter 2: Non-Compete Clauses and R&D Inventor Team Size

Chapter 1: Innovation in a Networked World: Inventor Teams, Social Networks and Patent Impact

Yanmin Yang

Department of Economics

Syracuse University

1. Introduction

One of the most striking trends characterizing the innovation of new ideas has been the declining presence of single inventors among all patent holders. While individual inventors developed the majority of patents in 1980, this ratio fell to roughly 30% by 2000. Concurrent with the decline of the individual inventor, large teams with at least four inventors increasingly are responsible for the generation of new patents (Figure 1). Shifts in firm size do not offer an easy explanation of this phenomenon: both large and small firms experienced notable increases in team size among patent inventors.[1]

Besides working more frequently in teams, patent inventors increasingly interact with inventors outside of their team. Figure 2 illustrates the sharp increase in co-inventor network size between 1980 and 2000, especially in large firms. In 2000, the average collaborative network size per patent was ten times larger than in 1980. The growth in inventor networks is consistent with Sorensen's (2018) argument that innovation is no longer an individual exercise but occurs in a "networked world". Social networks not only affect transmission and adoption of new knowledge, but also contribute to the impact of innovations (Sorensen, 2018).

Motivated by the two stylized facts above, this paper examines the role of team size and collaborative networks in knowledge creation. In particular, it determines whether social networks and teams are substitutes or complements in the production of knowledge. If the marginal effect of social networks on patent impact diminishes as team size expands, team size and social networks are substitutes. Alternatively, if larger team sizes amplify the impact of social networks, team size and social networks are complements.

[1] Small firms refer to firms with fewer than 100 patent applications from 1980 to 2000 and large firms refer to firms with more than 100 applications during the same period.

Using patent level data from USPTO during 1975 – 2000, my paper studies the joint effect of inventor networks and team size on patent impact and the patent development process. I focus on two outcomes of particular interest. The first, and the core of the paper, examines the impact of inventor networks and team size on the frequency with which a patent is cited. Patent citation counts are used as a measure of patent impact. The second outcome of interest is the ability of the patent team to access and build on newly created knowledge. This is measured both by the number of patents cited in the patent application and also by the inverse age of the latest citation which proxies for rapid access to new technology and knowledge.

For both outcome measures above, I must also address an identification challenge. It is likely that talented inventors join and/or have larger teams built around them. Therefore, a positive relationship between team size and patent impact, amount of knowledge inflow, or speed of knowledge adoption, does not necessarily indicate a causal effect of team size. The same concern is relevant when considering the relationship between inventor network size and the outcome measures above. To address this issue, I include interactions between network size and team size in the empirical model. The interaction term measures the marginal contribution of network size controlling for team size. It is expected to decrease (increase) as team size expands if inventor networks and teams are substitutes (complements).

Estimates from both outcome measures yield robust and negative coefficients on the interaction terms. This finding suggests that inventor network and team both provide access to knowledge and at least in part act as substitutes in the development of new patents. This result is especially strong when inventor networks and the patent under development are in the same technology class. It is also stronger for technology classes for which knowledge evolves rapidly, in contrast to less dynamic areas.

Through exploring the trade-off between inventor networks and team size, this paper makes four contributions to the innovation and social network literature. First, it examines the joint effect of inventor networks and team size on patent impact. While previous studies have considered the importance of team size in knowledge production (Wuchty et al., 2007; Jones, 2009; Singh and Fleming, 2010), they largely neglected the role of social networks. To my knowledge, this is the first empirical study of the trade-off between collaborative networks and team sizes on patent impact. Incorporating both networks and team size into the same model allows me to determine the substitutability between them empirically.

Second, my paper differs from the previous social network literature by focusing on patent quality instead of patent quantity. It is widely acknowledged that inventor network is important for knowledge transmission (Singh, 2005; Agrawal et al., 2008; Azoulay et al., 2010). However, there is no consensus on the role of inventor networks on knowledge creations (Lobo and Strumsky, 2008; Singh and Fleming, 2010; Breschi and Lenzi, 2016). Despite the mixed evidence on patent quantity, the effect of inventor networks on patent quality has received, surprisingly, much less attention. Patent quality cannot be measured using simple patent counts and its distribution is highly skewed (Trajtenburg, 1990; Harhoff et al., 1999; Hall et al., 2007). If inventor networks could boost both patent quantity and patent quality, the total economic impact of patents would be underestimated if patent quantity alone were evaluated.

Third, both inventor networks and teams are found to be important sources of knowledge input based on which new patent is developed. This finding contributes to the literature on the micro-foundations of knowledge spillovers and learning mechanisms (Rosenthal and Strange, 2001; Duranton and Puga, 2004; Azoulay et al., 2010). My paper proposes a new channel to explain the widely observed fact of an increasing dominance of large teams in developing high-impact patents (Wuchty et al., 2007). More specifically, larger teams are more likely to incorporate a broader source of knowledge into the patent development process, which ultimately improves the patent impact.

Fourth, this paper sheds light on the optimal rule of team formations and allocations in innovative activities, especially for entrepreneurs. The tradition model of team formation focused on the trade-off between specialization benefits and coordination costs within the team (Becker and Murphy, 1992). I extend this idea and take into account the extra benefit associated with the expansion of inventor networks. Therefore, the optimal rule to form inventor teams should also be dependent on the network size of each inventor. If inventor network and team are substitutes, the optimal team size could be smaller as inventor network size grows.

The remainder of the paper is organized as follows. Section 2 introduces how the network size measure in this paper is constructed. Section 3 describes the data. Empirical model is introduced in Section 4 and the corresponding regression results are reported and discussed in Section 5. Section 6 concludes the paper.

2. Background: network size measure

This section will introduce how network size, the key variable in the paper, is constructed using individual patent level data. I will start by providing some background in social network theory. Social distance are introduced afterward and used to construct the network size measure.

In social network theory, the relationship structure is usually illustrated in a tree graph like Figure 3. Individuals are denoted by nodes and social connections connect two nodes (individuals) with a line. Among many different types of social connections, this paper only focuses on connections formed through collaborations in patent development. In other words, two inventors would be directly connected if they were listed as inventors of the same patent.

Figure 3 illustrates the collaborative network for patent 001. There are six patents in Figure 3: patents 003 and 004 are developed by three inventors, and the remaining four patents have two inventors. The filing year of each patent is listed in the parentheses, and inventor identities are denoted by capital letters A through G.

Following prior research, social distance is defined as the geodesic distance between two inventors (Singh, 2005; Breschi and Lenzi, 2016). In Figure 3, the social distance between inventors A and B is one since they collaborated on patent 001. The social distance between inventors A and C is two as they share the same collaborator (inventor B) but have not worked together. The social distance between two inventors is positive if they are directly or indirectly connected and it is infinity if they are not connected.

I extend this original concept of social distance and construct a similar measure to calculate the distance between a patent and an inventor outside of the patent inventor team. For each patent, it is defined as the outsider inventor's minimum social distance to any inventor in the inventor team. This extension allows me to calculate network size at patent level afterward. Under this measure, inventor C is one unit of distance away from patent 001 although the social distance between inventor C and inventor A is still two.

For each patent, the relevant inventor network should include all inventors who contributed to patent development, excluding patent's own inventors. However, it is usually difficult to identify these contributors, as they are unobservable in my research data set. I use two criteria to define potential contributors for each patent. In general, these inventors should have a close social distance with any inventor on the patent inventor team during patent development. The following paragraphs explains the two criteria in details.

An inventor outside of the patent inventor team is more likely to contribute to patent development if this inventor has a short distance to the patent. In this study, the collaborative network is restricted to include inventors within two units of distance from the patent. This implicitly assumes that the relevant network members are past collaborators and their co-inventors. Restricting the distance to be within two units should not be a great concern. Although prior researchers observed knowledge transmission at a longer range, the magnitude of knowledge spillover diminishes as social distance increases (Singh, 2005).

The likelihood of contributing to patent development also increases if an inventor outside of the patent inventor team has worked with any inventors on the team during the production process of patent. In Figure 3, inventor G developed patent 004 with inventors A and B in 1988. It remains unclear whether inventor G had established relationships with the other two inventors, and contributed to patent 001's development back in 1985. On the other hand, the connection between collaborators might also depreciate over time. A past collaborator who has worked only once with the inventor ten years ago is very unlikely to contribute to the inventor's current patent development. Hence, I further restrict the collaborative network to include inventors who has worked with any inventors on the patent inventor team within two years of the patent's filing year, excluding patent's own inventors.[2] In Figure 3, the valid co-inventor network for patent 001 includes only inventors with patents filed between 1984 and 1985. Inventors F and G are not included because their collaborations are either too new or too old.

Therefore, the network size measure used in this paper is defined as the total number of inventors in such a network that satisfies the two conditions described above. To be specific, the

[2] Since I cannot observe the working experience of inventors, two inventors are considered to have worked together if they are both listed as inventors of the same patent. The filing year of the corresponding patent marks the year when they collaborated.

measure counts the total number of inventors within two years of patent's filing year, who are also no more than two units of distance away from the patent in question. The own inventors of the patent are excluded. This is similar to the network size measure used in Singh and Fleming (2010). It captures both size and connectivity of the collaborative network. It is also computationally much easier than calculating the size of the complete co-inventor network. In Figure 3, the corresponding network size for patent 001 is three (inventors C, D and E).

There might be concerns due to lags between the timing of collaborations and filing of patent applications: inventors must have worked together *before* patents are filed. For example, patent inventors could have consulted inventors outside of their team in 1984 for a patent filed in 1985. The original measure with the two-year window before the filing year of patents would fail to capture such collaborations. To address the concern, I constructed a similar measure with a three-year window. The new network further includes other inventors who worked with inventors on the patent inventor team one year *after* the filing year of the patent. Results using this new measure are reported in Section 5.3.

This paper further separates the network into two types according to inventors' specialized field. Prior studies implicitly treated network members as homogeneous (Singh, 2005; Singh and Fleming, 2010). However, inventors from different technology fields can contribute to patent development in different ways. It is well known that innovations usually require recombination of existing knowledge from different areas. However, the "burden of knowledge" (Jones, 2009) makes it difficult for inventors to learn and acquire knowledge across fields. While inventors with different knowledge backgrounds might contribute greatly to patent development, the benefit could be limited if inventors share similar knowledge backgrounds with current team members.

Since I cannot observe the specialized field of each inventor with the current data set, I use the main patent classes from patents developed by inventors as their specialized fields. The full network is then split into two according to the assigned specialized fields. For each patent, one network includes all inventors sharing the same patent class as the patent, and the other one includes all other inventors. It is very likely that inventors in the first network share a similar knowledge background as inventors in the patent inventor team. To illustrate the two networks in Figure 3, I use black circles and black squares to denote two separate technology fields. Therefore, the size of the network sharing the same technology field as patent 001 is one (inventor C) and the size of the other network is two (inventors D and E).

3. Data and variables

The main research dataset contains all patents granted by the United States Patent and Trademark Office (USPTO) from 1975 to 2010.[3] It provides detailed information on each patent: name and residential address of each inventor, patent class, patent filing year, citations made (*input citations*) and received by the patent (*output citations*), etc. Since the original data set lacks a good measure of a unique patent assignee (firm) identifier, I complemented it with data from NBER patent project.[4] I also matched each patent to a broad technology category as defined by Hall et al. (2001).

This paper focuses on utility patents assigned to US corporations with at least one inventor in the contiguous US. This limits the inventor connections to be completely within the contiguous US. Each patent is assigned to one core-based statistical area (CBSA) corresponding

[3] This dataset comes from Li et al. (2014).

[4] This data can be downloaded from https://sites.google.com/site/patentdataproject/Home. If there are multiple assignees, the first one is used.

to the location of the majority of inventors.[5] Patents with all inventors from micropolitan statistical areas are dropped, which accounts for around 6% of the data set.

The sample period ranges from 1980 to 2000. This leaves enough time for follow-on citations, which are future patents, to show up. Excluding samples in the first five years allows me to construct prior patenting history to control for inventor team heterogeneities. This paper does not follow Singh (2005), which drops patents that are not matched to Compustat database. This process involves a non-random sample selection: observations from small firms with no data available in Compustat are likely to be dropped.

The main dependent variable is the logarithm of number of citations the patent received by 2010.[6] This measure is found to be closely correlated with the market value of firms (Harhoff et al., 1999; Hall et al., 2005; Hall et al., 2007; Kogan et al., 2017) and it is commonly used as a proxy for patent quality (Trajtenberg, 1990; Jaffe et al., 1993; Harhoff et al., 2003; Hall et al., 2005; Kerr, 2010). It also allows me to evaluate patent impact across different technology fields in a systematic way.

However, citations are not random. Patent inventors might cite patents because of cronyism, which is not necessarily related to patent impact. Therefore, I only use number of citations received, after excluding self-citations, as the patent impact measure.[7] The distribution graph of citation counts is illustrated in Figure 4. Robustness checks with different patent impact measures are reported in Section 5.3.

Two more dependent variables are constructed to describe the process of patent development. By extracting detailed information from citations listed on the patent application, I

[5] This is similar to Kerr (2010) and I assign the address according to the sequence of authors if there are ties.
[6] To be specific, the dependent variable is ln(1 + number of citations received). I add one before taking logarithm because some patents received no citations by 2010.
[7] Self-citations include citations made by the same inventor or from the same assignee (firm).

calculated number of citations and age of the latest citation. The first measure is used to capture the amount of knowledge inflow. I take the inverse of the second measure to capture the speed of new knowledge adoption: adopting new knowledge occurs more quickly when age of the latest citation is smaller. In the robustness check, I also calculated the mean age of citations and used its reciprocal to proxy for the speed of knowledge adoption. Age of each citation is calculated in days: it is the gap between filing dates of the citation patent and the cited one. All dependent variables are taken logarithm in the analysis.[8]

The construction of the main independent variable, network size, requires that I follow individual inventors across patents. However, the USPTO does not have a unique identifier assigned to the same inventor. This paper uses the inventor identifier generated by Li et al. (2014) to construct the co-inventor network. For each patent, the collaborative network is further split into two according to inventor's specialized field. One network includes all inventors sharing the same technology class as the patent, and the other one includes all other inventors. I calculated the network size as described in Section 2 for all the three networks.

There might be concerns due to the process of matching inventors across patents. Li et al. (2014) generated the unique inventor identifier through their disambiguation algorithm. If this algorithm provides a good match for every single inventor, it would allow me to clearly identify who exactly collaborated with whom. However, inventors can still be erroneously matched. This should result in standard measurement errors as it is not clear in which direction network sizes would be affected.

Through focusing on the co-inventor network, I neglected other ways like conversations and meetings, through which knowledge is shared and exchanged between inventors. Firms

[8] I add one to all measures with a minimum value of zero before taking logarithm.

could also share technology when they form strategic alliances (Schilling and Phelps, 2007). However, these relationships are unobservable using the current data set. It is also uncommon to see patents assigned to multiple firms.

The other important independent variable is team size: it is defined as the total number of inventors for each patent. In this paper, I categorized the team size measure into four groups: one inventor, two inventors, three inventors, and at least four inventors. Only around 15% of patents are developed by large teams with at least four inventors (Table 1). I also constructed two patent-level control variables for patenting experience. For each patent, I aggregated the total number of patents developed by any team members in the last five years as well as the total number of patent classes on these patents. These two measures are also used in Singh and Fleming (2010). They capture the experience and diversity of previous patenting history.

In recent years, inventors not only collaborate with people sharing the same specialized field, but they also work with inventors who have different backgrounds. From Figures 5 and 6, I find that both networks expand in size in all broad technology categories, especially in "electrics and electronics" after 1995. This is consistent with the information technology boom in the 1990s.

Summary statistics for all other variables are reported in Table 1. All measures are reported in levels. Around 47% of patents have network size of zero and fewer than 5% of patents have network size above 50. Six patents were developed in last five years for the average inventor team. Inventors also specialize in a very narrow range of defined technologies: the average number of patent classes for their teams in the last five years is around two.

4. Empirical model and hypotheses

4.1 Regression model

This is an individual patent level study. Considering patent i developed in a CBSA in year y in patent class c, the main regression model is as follows:

$$\begin{aligned}\text{outcome}_{\text{i,CBSA,y,c}} = & \beta_1 \log\left(\text{network size}_{\text{i,CBSA,y,c}}\right) + \beta_2\left(\text{Team size}_{\text{i,CBSA,y,c}}\right) \\ & + \gamma \log\left(\text{network size}_{\text{i,CBSA,y,c}}\right) \times \text{Team size}_{\text{i,CBSA,y,c}} \\ & + \tau \text{X}_{\text{i,CBSA,y,c}} + \eta_{\text{CBSA}} + \xi_{\text{y}} + \theta_{\text{c}} + \varepsilon_{\text{i,CBSA,y,c}} \qquad (1)\end{aligned}$$

My paper mainly focuses on the effects of network size, team size and their interaction terms on the main outcome, patent impact. In equation (1), γ is the parameter of interest: it captures the marginal effect of network size expansions for large teams versus individual inventors. If this coefficient is negative, it means that the marginal effect of network sizes for large teams is smaller compared to the base level (a single inventor). In other words, patent team size and network size are substitutes in determining patent impact. In contrast, a positive coefficient of γ indicates that team size and network size are complements.

$\text{X}_{\text{i,CBSA,y,c}}$ is the set of control variables for patent level characteristics. With this paper's construction of network size measure, it is obvious that patents with network size above zero must include experienced inventors who have longer patenting histories. Therefore, I included patenting experience control variables as described in Section 3. I also included year fixed effect, CBSA fixed effect and patent class fixed effect to control for heterogeneities across periods, regions and industries.

4.2 Testable hypotheses

In this section, I will introduce three channels through which patent inventor team size and co-inventor network size jointly affect patent impact. In particular, I will discuss predictions of the sign of γ for each channel respectively. Since γ is the coefficient of the interaction term in equation (1), its sign infers whether a large co-inventor network size amplify or erode the effect of a large team size on patent impact.

γ could capture a substitution effect in the patent production process. Teams and social connections are viewed as important channels through which new knowledge for patent development is transferred and acquired. If knowledge learned through these two channels are overlapping, teams and co-inventor networks can act as substitutes.

Hypothesis 1: *γ is negative through the channel of patent production function. The negative γ represents the substitution effect between teams and social networks.*

However, it is difficult to rule out sorting of patent inventors based on unobserved abilities. Observing positive values of β_1 and β_2 might not imply an intrinsic advantage of large teams or social networks in patent development. It could simply represent sorting. Talented inventors are more likely to develop high-impact patents. They are also more likely to have large professional networks and work in large teams.

If inventor's ability level is positively correlated with the inventor team size and network size, γ is expected to be positive. High-ability inventors are more likely to get involved in patent development if both a large team and network are observed. Hypothesis 2 summarizes the prediction:

Hypothesis 2: *γ is positive through the sorting channel. The positive γ represents the fact that the likelihood of having high-ability inventors in the patent development process is greater when both large teams and large network sizes are observed.*

Since patent impact is measured by number of citations received in this paper, there are naturally concerns of cronyism: citations might demonstrate friendship rather than patent impact. In other words, it is likely that patents developed by large teams and associated with heavy social connections would receive more citations simply because their inventors have more friends and friends of friends. However, ideas are expected to spread even further as the chain of friendship extends. Therefore,

Hypothesis 3: γ *is positive through the cronyism channel. Having both a large team size and network size is expected to create more citations through friends of friends.*

The empirical study captures effects from all three channels. Therefore, the coefficients of γ in the empirical results would represent the aggregate effects. Among these three hypotheses, the first one is the only one through which γ is predicted to be negative. Hence, if a negative γ were observed, it is likely to be the lower bound of the substitution effect.

5. Empirical results

5.1 Network size, team size and patent impact

5.1.1 Basic results

Table 2 presents the results of the joint effect of co-inventor networks and team size on patent impact. CBSA fixed effect, filing year fixed effect and patent class fixed effect are controlled in all regressions. Errors are clustered at firm level. The dependent variable for all regressions is the logarithm of number of citations received by 2010. The first column reports the basic result. The last two columns display results with interaction terms, and the last column further includes controls for past patenting experience.

Column 1 shows that both expansions of network size and team size increase patent impact. If network size expands by 1%, number of output citations is likely to increase by 1%. The coefficients of team size indicators are significantly greater than zero and robust across three columns. It is also worth noting that these coefficients grow monotonically as team size increases. Large teams are more likely to develop high impact patents, as compared to individual inventors, *ceteris paribus*. These findings are consistent with Wuchty et al. (2007) and Singh and Fleming (2010).

Table 2 also provides evidence that the marginal effect of network size on patent impact decreases as team size expands. The interaction terms are significantly negative in both columns 2 and 3. After including the interaction terms in column 2, the network effect for solo inventors increases to 2%. Meanwhile, the marginal effect of network size on patent impact reduces to 1% for large teams with at least four inventors. This result does not change much after including patenting experience control variables in column 3.

The previous findings should be interpreted with caution. According to Hypotheses 2 and 3, sorting on unobserved ability and cronyism could also lead to positive values of β_1 and β_2. However, it would be difficult to interpret the findings in Table 2 if these mechanisms were the only confounding factors. Instead, the negative coefficients of the interactions terms can be better explained by Hypothesis 1: team size and network size are substitutes in the patent production process.

As team size grows, the interaction terms get smaller but the differences across these coefficients are not significant (columns 2 and 3). There are two potential explanations. First, the aggregate size of co-inventor networks is too broad a measure and adds noise to the result. Different network members provide different knowledge for patent development. The substitution effect is more likely to come from a particular group of inventors in the network who shares a similar knowledge background as patent inventor. This channel will be checked in Section 5.1.2. Second, this could also come from the sorting mechanism as described in Hypothesis 2. The sorting mechanism could bias the coefficients of the interaction terms towards zero, if not positive. If sorting is stronger in larger teams, it is also possible that all interaction terms would have similar coefficients.

In Table 3, I study the joint effect of team size and network size on patent impact distribution. Instead of focusing on citation counts, the regressions in Table 3 examine the likelihood that a patent is high-impact or low-impact. Patents with top 10% citations within the same year and patent class cohort are considered as high-impact ones. In contrast, low-impact patents refer to those without citations by 2010. Columns 1 and 2 report empirical results of the probability of a patent being a high-impact one and a low-impact one, respectively.

The coefficients on team size and network size have expected positive signs. However, the interaction terms at the two ends of the distribution are slightly different. Team size and social network size are found to be substitutes in developing high-impact patents, but there is no significant interaction between them in avoiding poor outcomes.

5.1.2 Networks separated by specialized fields

As is mentioned above, the aggregate network size measure might have included many heterogeneous types of inventors. This section will separate the full network into two different types of co-inventor networks as described in Section 2. *NS_SC* is used to label the size of the network that shares the same technology class with the patent, and *NS_DC* is used to label the other one. Regression results with both network size measures are reported in Table 4.

Columns 1 and 2 report results when only one network is included. In columns 1 and 2, I find that two network size measures increase patent impact individually. However, the substitution effect is only significant for the network sharing the same patent class with patents (*NS_SC*). Column 3 reports results when both measures are included. From column 3, a 1% increase in NS_SC would increase patent citations by 1.4% for solo inventors. However, for large teams, the marginal effect is negative: it reduces by 1.8%. It is also worth noting that the interaction terms for NS_DC are not significant and much smaller in magnitude in column 3. This could infer a small substitution effect or a strong sorting effect that bias the results towards zero.

Compared with previous findings in Table 2, the substitution effect between inventor team and one particular co-inventor network is much stronger and more significant in Table 4. Since the interaction terms in Table 2 represent the aggregated results of two networks, the results in Table 4 should not be unexpected. It implies that there is a strong substitution effect between inventor teams and networks when they share the same patent class. For network members with different patent classes, they could provide different information for patent development and do not act as substitutes. However, I could not rule out the possibility that there are more erroneous matchings based on inventor names, especially for inventors with patents across multiple technology classes. This could also lead to insignificant coefficients of interactions terms for NS_DC.

Results from Table 4 suggest that the substitution effect could be explained by the overlapping knowledge base because team members and social network members within a narrowly defined patent class are very likely to have similar knowledge backgrounds. I will discuss that and provide related empirical evidence in Section 5.2.

5.1.3 Fast-evolving industries versus slow-evolving industries

The substitution effect is not necessarily the same across different technology fields. It depends on how knowledge is learned, acquired and adopted. In this section, I separate the sample into two groups according to the reliance of new technology and knowledge in the industry (patent class). In each year, I calculated the mean age of citations for each patent class.[9] Patent classes in the bottom 50 percentile of average citation age are defined as *fast-evolving* ones since most of their inputs are newer knowledge. Similarly, patent classes with older citations are labeled as *slow-evolving* ones. It is calculated for each year since the technology trend could be different across times.

Table 5 reports the regression results for both fast and slow evolving patent classes. Significant and negative interaction terms can be only observed in fast-evolving patent classes in column 1. This suggests potential differences in patent production processes across different industries. For patent classes that are built on very recent knowledge, it is very likely that network members share similar and limited knowledge base with team members. This limits the benefit from expanding the inventor network size. I also conduct similar analysis as in Table 3 and check the likelihood of a patent being a high-impact or low-impact one for both types of patent classes. Corresponding results in Table A1 are consistent with previous findings.

Results are still consistent with previous ones when the network size measure is split into two based on whether network members and the patent share a same patent class. Table 6 reports the results for different subsamples with both network size measures. I can only observe a strong and significant substitution effect in fast-evolving industries. The substitution effect is more significant when inventor teams and networks share the same patent class.

5.1.4 Including firm fixed effect

Empirical results of previous tables rely on variations of inventor team size and network size within the same CBSA and patent class. It is thus hard to rule out sorting across firms in the same area. Firms with longer patenting history and larger number of patent applications are more

[9] Patent classes with fewer than 10 patents are dropped.

likely to have experienced managers for inventor team assignment. They could also have better institutions and organizational structures, both of which could affect efficient team formation and patent impact (Bradley et al., 2016). If these firm-specific factors are time-invariant, I could reduce the concern of the sorting problem by including firm fixed effect.

Table 7 reports results with firm fixed effects. The identification for these regression models comes from within firm variations. Again, the coefficients of interactions terms for NS_SC are highly significant and negative, only in fast-evolving patent classes. In the appendix Tables A2 and A3, I further report results for stratified subsamples in fast and slow evolving patent classes separately. Results for both high-patenting and low-patenting firms are similar to findings in Table 7.

In summary, the robust findings of negative coefficients of interaction terms suggest a substitution effect between teams and inventor networks in determining patent impact. It mainly takes place in fields where teams and networks share the same knowledge background. This effect is more dominant in industries that largely rely on new knowledge. These findings are not greatly affected after controlling for firm fixed effects.

I also followed Agrawal et al. (2008) to quantify the substitution effect between team sizes and network sizes. I estimated the model with only team size indicators, logarithm of network sizes and other fixed effects (e.g. without interactions). [10] The coefficients on three team size indicators are 0.050, 0.106, 0.200 and the coefficient on network size is 0.022. These results imply that the net effect of increasing team size by one is equivalent to expanding the network size by 8.71 people.[11]

[10] CBSA FE, patent application year FE, patent class FE and firm FE are included.
[11] This estimate is obtained by forming exp (0.050/0.022) -1 $\approx$ 8.71.

5.2 Network size, team size and knowledge adoption: a potential explanation

Findings in Section 5.1 suggest that patent inventor teams and networks are substitutes in determining patent impact. According to Hypothesis 1, this is rooted in the patent production process: teams and networks act as substitutes in adopting new knowledge for patent development, and this ultimately determines patent impact. In this section, I will provide direct empirical evidence that links Hypothesis 1 with findings in Section 5.1.

It is usually difficult to measure which knowledge is applied to patents, and where this specific piece of information is learned. In this paper, I use information from citations to measure knowledge adoption. However, I am not able to distinguish whether citations are added by examiners or inventors during the sample period. [12] Therefore, all citations in the dataset, excluding self-citations, are considered.

5.2.1 Amount of knowledge adopted

Table 8 reports empirical results of inventor networks and team size on the logarithm of number of citations, which proxies for the amount of knowledge used for developing a new patent. The correlation between number of input citations and number of output citations is around 0.12 for the whole sample.

For the fast-evolving patent classes in column 1, I find that the expansion of team sizes increases the amount of knowledge input. Compared to solo inventors, large teams with at least four inventors have 14% more input citations. This suggests that large teams usually have a broader source of knowledge. The coefficient of NS_SC for solo inventors is negative but it is insignificant. On the other hand, the network with different technology classes from the patent being developed still plays an important role. For 1% increase in NS_DC, the number of

[12] I can only observe examiner added citations after 2001.

citations increases by 1% for patents from fast-evolving patent classes and 3% for those from slow-evolving ones.

The coefficients of the interaction terms between NS_SC and team size indicators in column 1 are negative: the marginal effect of this particular network on the amount of knowledge inflow diminishes as team size increases. This finding does not hold for slow-evolving patent classes in column 2. These findings are consistent with previous findings in Section 5.1 and provide direct support for Hypothesis 1.

5.2.2 Speed of knowledge adoption

This section reports the joint effect of team size and network size on the speed with which new knowledge is adopted. In Table 9, the dependent variable is the logarithm of the reciprocal of age the latest input citation. There are several findings worth mentioning. First, knowledge adoption occurs more quickly in larger teams. Compared to solo inventors, the speed of knowledge adoption is 10% faster for teams with at least four inventors (columns 1 and 2). Second, the coefficients on both network size measures are positive in columns 1 and 2: patents associated with larger social networks are more likely to rely on knowledge inputs from recently developed patents. This result holds for patents in both fast-evolving industries and slow-evolving industries.

Furthermore, when considering the speed of knowledge adoption, the marginal effect of network size based on connections within the patent's classification diminishes as inventor team size increases. This is especially true for fast-evolving patent classes. This suggests that inventor teams and inventor networks that belong to the same patent class are substitutes for purposes of learning about and adoption of newly developed knowledge. While a larger team or a larger social network would facilitate knowledge adoption, the net benefit from the network is smaller as team size expands. These findings provide extra supporting evidence for Hypothesis 1.

Concerns might arise because citations are sometimes added by examiners, which does not represent knowledge adoption by inventors. I believe this is not likely to affect my estimates of the speed of knowledge adoption for the following reasons. If the age distribution of patent citations added by examiners is similar to patents cited by the inventor team, then examiner added citations will not affect estimates of the speed of knowledge adoption. On the other hand, if examiners disproportionately add recently developed patents to patent applications, the age of the latest citation should not vary much across patents with different inventor team sizes. This is in contrast with the findings in Table 9, where the age of the latest input citation is smaller as team size increases. [13]

As a robustness check, I also use the reciprocal of the mean age of citations to measure the speed of knowledge adoption. Corresponding results in Table A5 are similar to those in Table 9. Therefore, I conclude that my findings here support my suggested theory and they are all consistent with Hypothesis 1.

5.3 Robustness

Four robustness checks are reported in this section. First, Table 10 includes regression results using different measures of patent impacts. In column 1, I use total number of citations received by 2010 as the patent impact measure, including self-citations. Although I excluded self-citations in the main analysis, it might be unnecessary. Hall et al. (2005) found that self-citations are "better measures" of patent quality. In column 2, I use number of citations, after

[13] Recall that the dependent variable in Table 9 is the logarithm of the inverse age of the latest citation. As team size increases, the age of the latest citation decreases because its reciprocal increases.

excluding self-citations and citations from the same CBSA as the patent impact measure. Since this paper is concerned with citations made due to friendship, this new measure can address the issue through ruling out citations made through local connections. The results in Table 10 are similar to previous findings. They are not largely affected after splitting the sample into fast-evolving and slow-evolving patent classes (Table A4).

Second, it is possible that my empirical results are driven by the network size measure constructed on narrowly defined patent classes. Different patent classes could still represent similar technology fields. In fact, many patents have multiple patent classes listed in the data set and only the major one is used in this paper. To address this concern, I split the network size measure into two based on connections within the patent's broad technology category, instead of the technology class. This broad technology category is a much wider measure and contains only six categories. The corresponding results reported in Table 11 are consistent with Hypothesis 1.

Third, my empirical results could be driven by computer-related industries. Using the classification from Hall et al. (2001), almost all patents categorized as "communications and computers" fall into the fast-evolving group. Therefore, I report several empirical results after excluding this technology category in Table 12. Although the direct effect of network size is not significantly different from zero in column 3, it is still positive in sign. The substitution effect remains in all regressions.

Finally, this paper's empirical results are also not driven by the two-year time window used to construct the network size measure. I constructed a new network size measure based on a three-year time window as described in Section 2. As a repeated exercise for Table 4, I report the separate results for both network size measures in Table 13. Table 14 is similar to Table 6, which further presents results in fast and slow evolving industries, respectively. [14] The empirical results are robust under the new measure.

6. Conclusions

This paper studies the joint effect of inventor network size and team size on patent impact and provides theory and supporting empirical evidence that networks and team size are substitutes in knowledge creation. To obtain this result, I separate the inventor network size into two types of measures and conduct several detailed analyses on patent impact. Moreover, I also study the joint effect of network and team size on knowledge adoption, which confirms that the previous finding can be explained by the substitution effect between these two factors as sources of knowledge inputs.

The main finding in this paper is a strong substitution effect between inventor networks and teams. For patents developed by individual inventors, considering the associated network that belongs to the same technology class, 1% increase in its size will increase patent impact by 2%. As team size increases, the effect on patent impact from this particular network reduces. It is concluded that team size and inventor networks are substitutes in determining patent impact. This is especially the case in fast-evolving industries and when patent inventor teams and network members share an overlapping knowledge background.

I also conduct empirical analyses on the patent development process. In particular, this paper examines the joint effect of inventor team size and network size on the amount of knowledge input and the speed of knowledge adoption. Both team size and networks are positively correlated with the amount of knowledge inflow and the speed of knowledge adoption. However, as team size grows, the marginal effect of network size on the amount of knowledge

[14] Year 2000 is dropped from the sample period.

inflow and the speed of knowledge adoption diminishes. These empirical results corroborate the hypothesis that inventor networks and teams are substitutes. They also provide a channel to explain previous findings on patent impact. Patent impact is likely to be determined in the patent development process based on knowledge learned through teamwork or through inventor's network. My findings of the substitution effect in patent impact could be explained by an overlapping knowledge base between inventors and networks in the production of a new patent.

This research provides new empirical evidence of inventor networks on innovations and proposes a new channel to better understand the micro-foundations in patent development. It also sheds light on the optimal rule of team formation in R&D. The optimal team size in R&D could be smaller if inventor networks and team size at least in part act as substitutes. It will be helpful for entrepreneurs who would like to invest in new R&D projects and develop new patents.

There are still some limitations in the study. Although I control for many types of fixed effects and include patenting experience measures, I am still not able to completely rule out the sorting problem of inventors into different teams. However, my empirical results provide at least a lower bound for the substitution effect. For future research, it will be interesting to discover whether this substitution effect can be avoided or minimized under certain network structures. I will further explore other determinants of patent impact and their relationships with team sizes in the networked world.

References

Agrawal, A., Kapur, D., & McHale, J. (2008). How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics*, 64(2), 258-269.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics,* 125(2), 549-589.

Becker, G. S., & Murphy, K. M. (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics, 107*(4), 1137-1160.

Bradley, D., Kim, I., & Tian, X. (2016). Do unions affect innovation?. *Management Science*, 63(7), 2251-2271.

Breschi, S., & Lenzi, C. (2016). Co-invention networks and inventive productivity in US cities. *Journal of Urban Economics*, 92, 66-75.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, 4, 2063-2117.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 36(1), 16-38.

Hall, B. H., Thoma, G., & Torrisi, S. (2007). The market value of patents and R&D: evidence from European firms. *Academy of Management Proceedings*, 1, 1-6.

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics*, 81(3), 511-515.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy*, 32(8), 1343-1363.

Sorenson, O. (2018). Innovation policy in a networked world. *Innovation Policy and the Economy, 18*(1), 53-77.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3), 577-598.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?. *The Review of Economic Studies*, 76(1), 283-317.

Kerr, W. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of*

Urban Economics, 67(1), 46-60.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, 132(2), 665-712.

Li, G. C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., Amy, Z. Y. & Fleming, L. (2014). Disambiguation and co-authorship networks of the US patent inventor database (1975–2010). *Research Policy*, 43(6), 941-955.

Lobo, J., & Strumsky, D. (2008). Metropolitan patenting, inventor agglomeration and social networks: A tale of two effects. *Journal of Urban Economics*, 63(3), 871-884.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics, 50*(2), 191-229.

Schilling, M. A., & Phelps, C. C. (2007). Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science*, *53*(7), 1113-1126.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*, *51*(5), 756-770.

Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality?. *Management Science*, *56*(1), 41-56.

Sorenson, O. (2018). Innovation policy in a networked world. *Innovation Policy and the Economy*, *18*(1), 53-77.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, 172-187.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036-1039.

Figure 1: Distribution of Inventor Team Size by Year

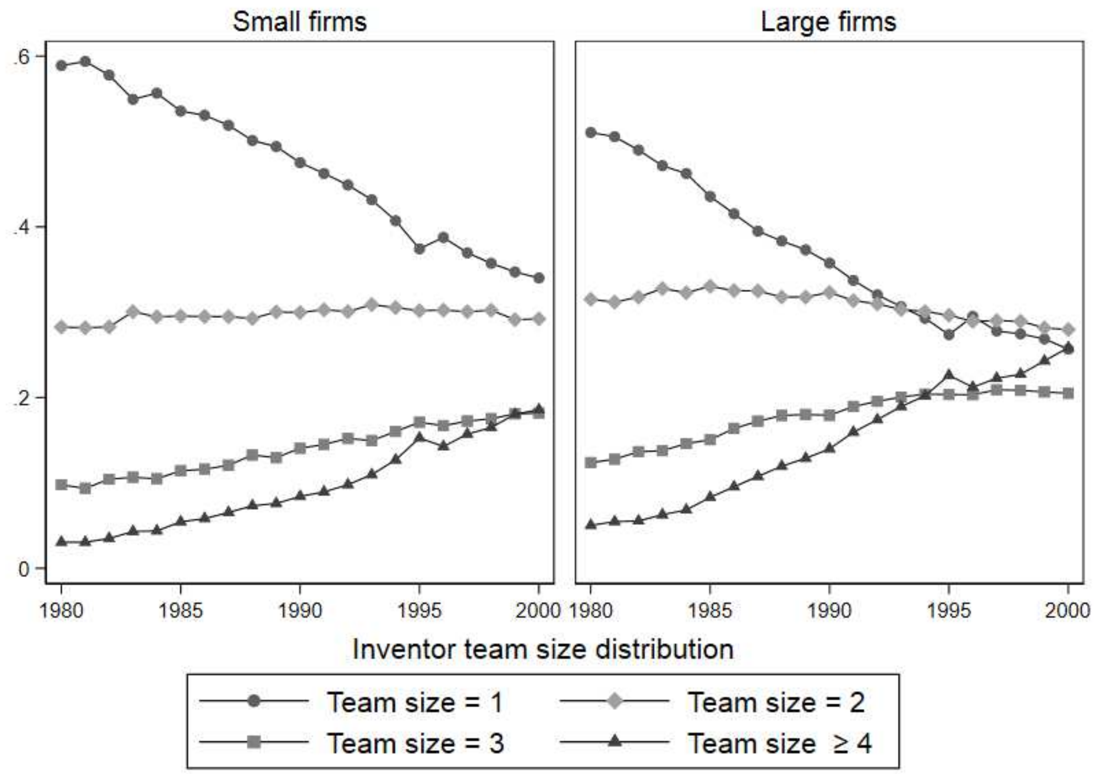


Figure 2: Mean Size of Inventors' Collaborative Network

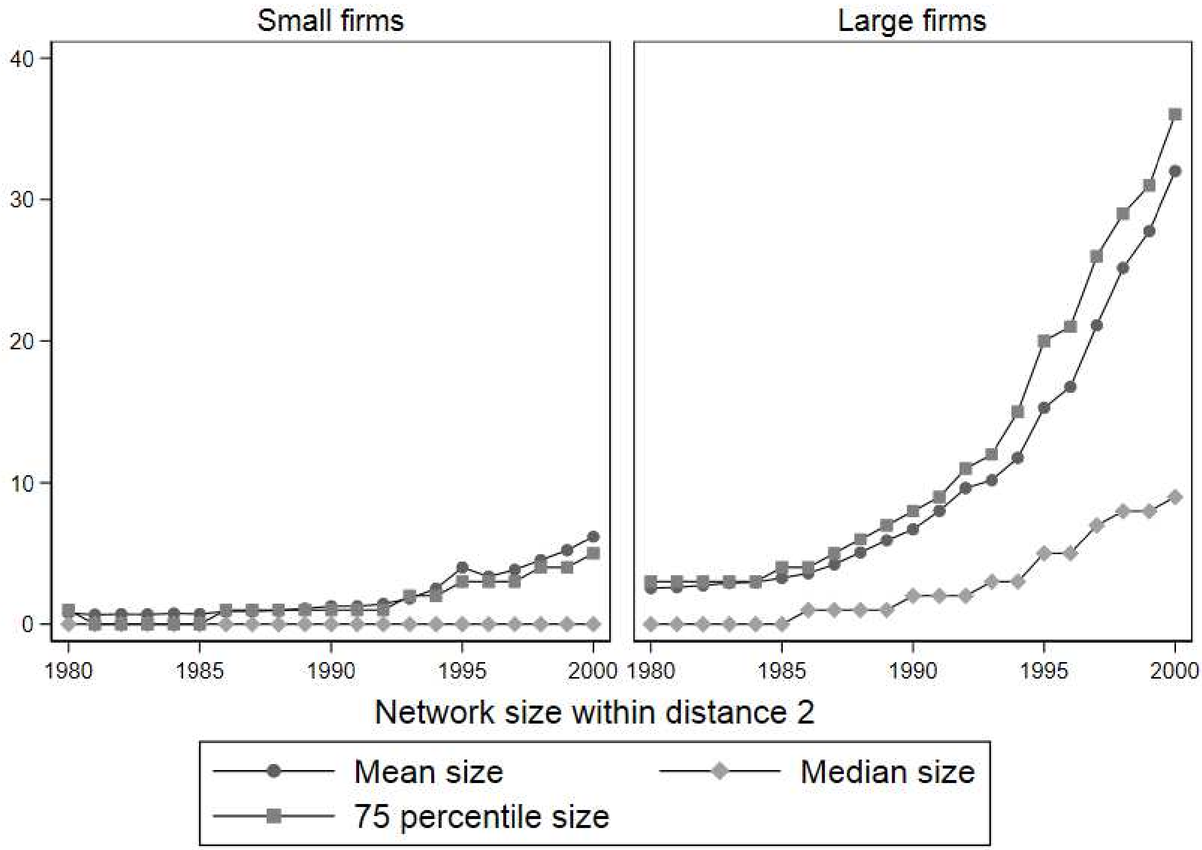


Figure 3: Collaborative Network for Patent 001

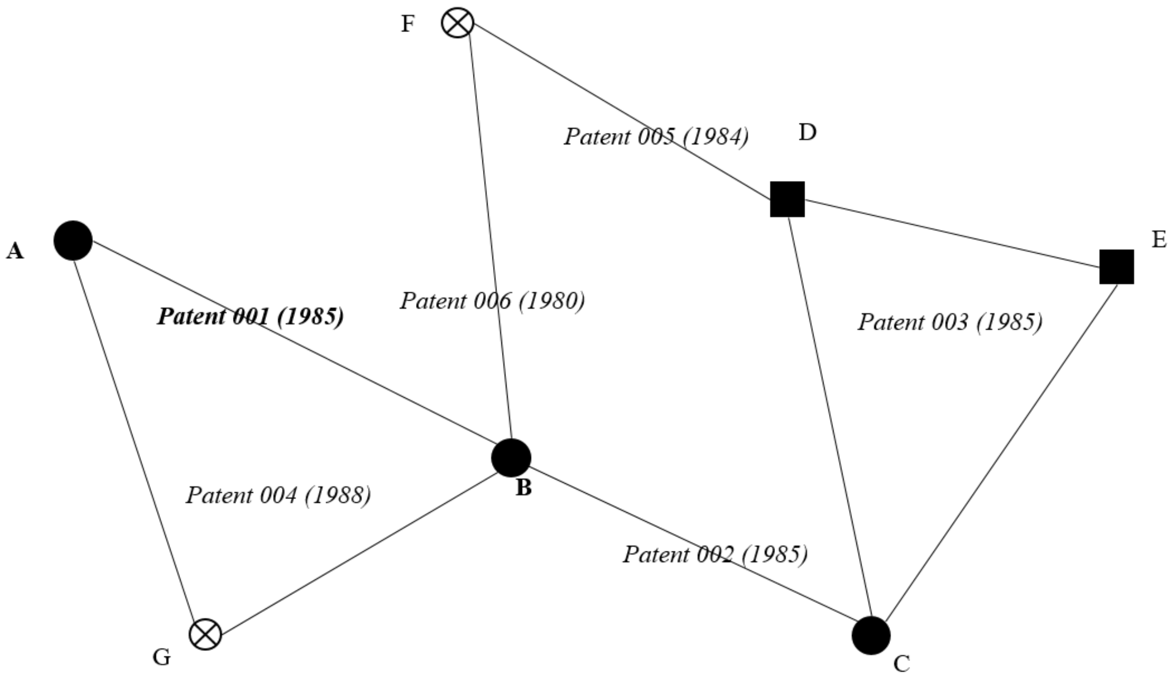


Figure 4: Distribution of Number of Citations Received by 2010 (Truncated at 100)

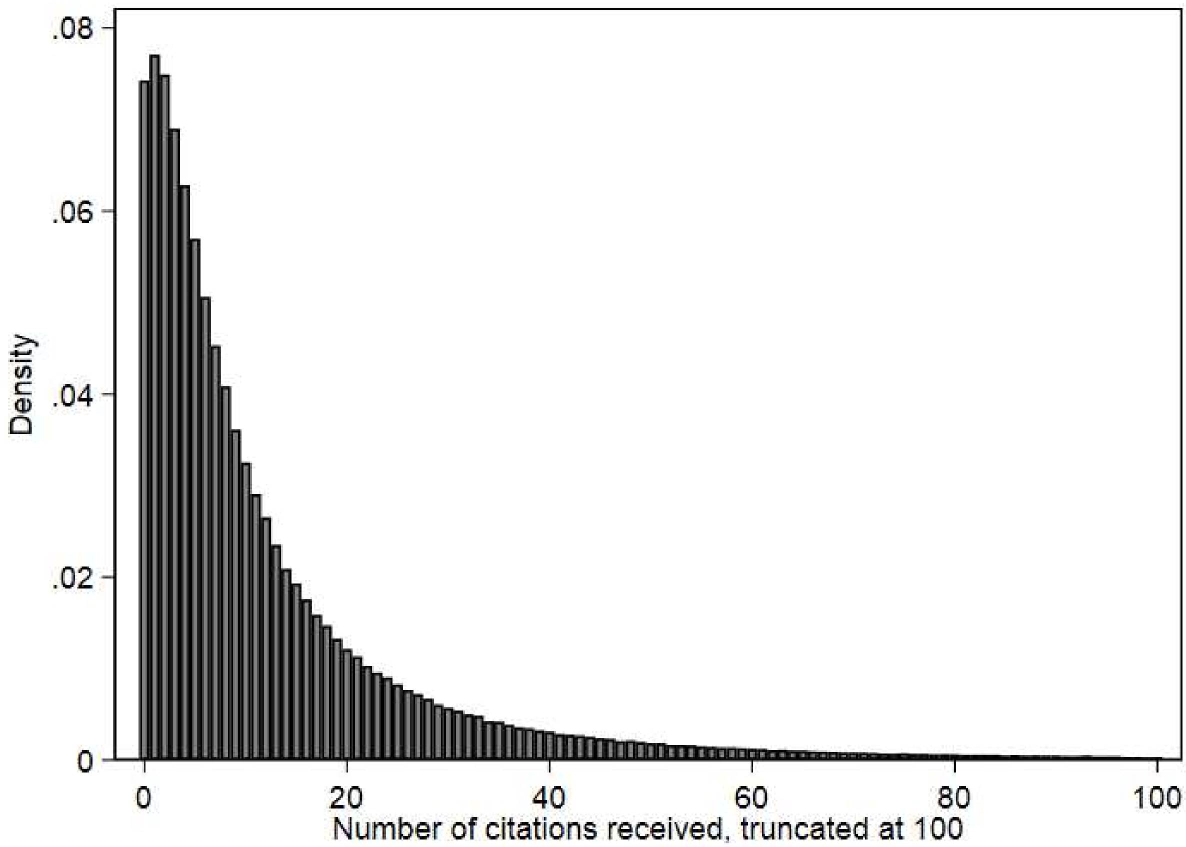


Figure 5: Mean Size of Collaborative Network with the Same Patent Class as Patent

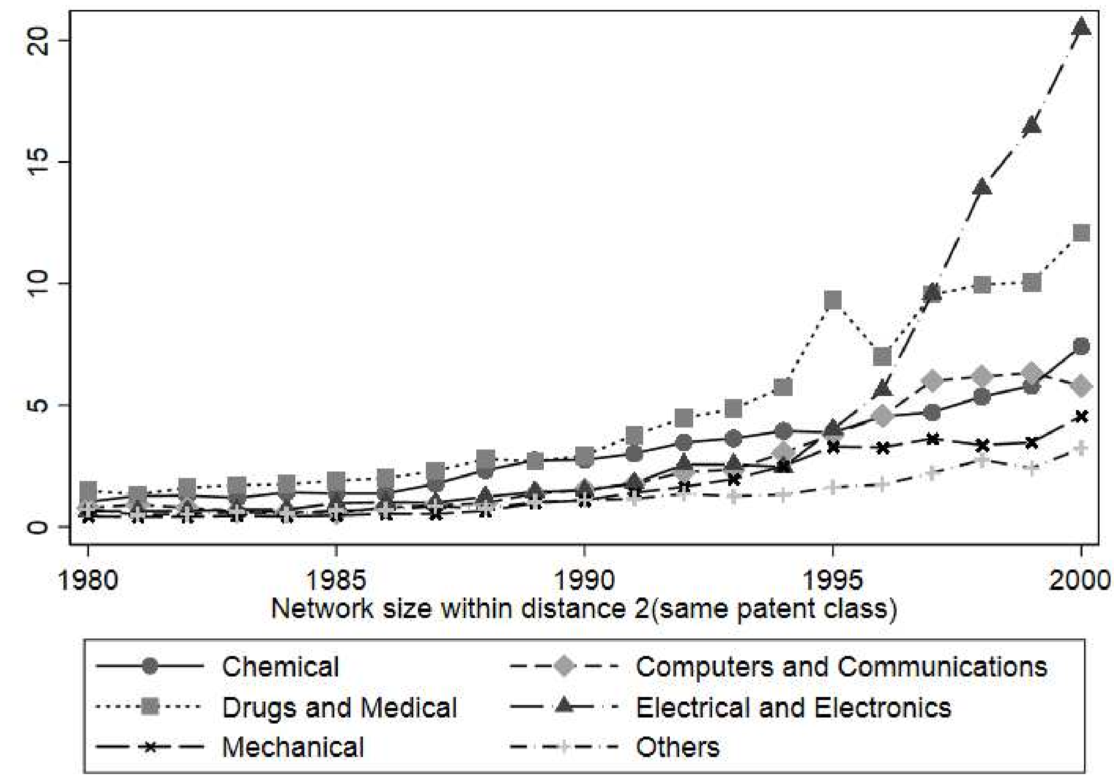


Figure 6: Mean Size of Collaborative Network with Different Patent Classes from Patent

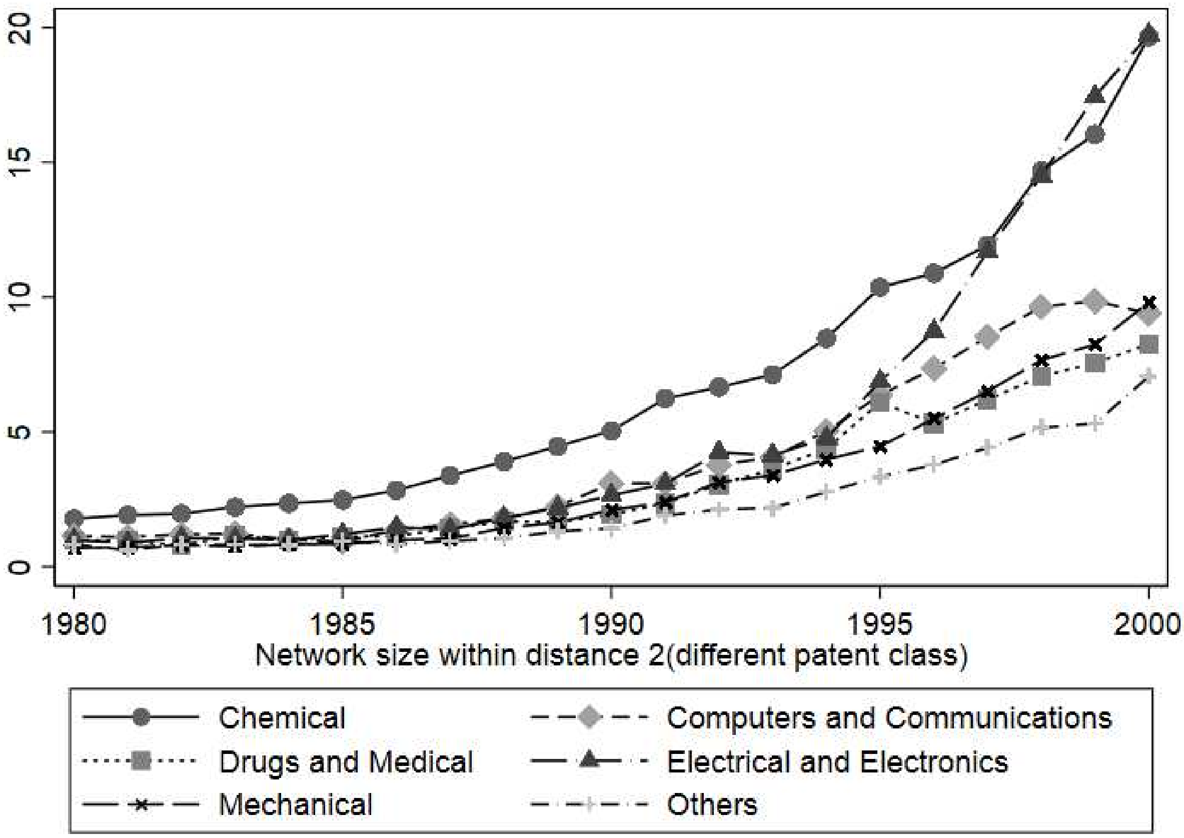


Table 1: Summary Statistics

Variable	Obs.		Mean	Std. Dev.	Min.	Max.
Patents by number of inventor	868,825		2.25	1.50	1	41
1 inventor	324,943	(37.4%)				
2 inventors	261,559	(30.1%)				
3 inventors	150,497	(17.3%)				
At least 4 inventors	131,826	(15.2%)	5.00	1.63	4	41
Network size within distance 2 within last 2 years	868,825		9.71	27.51	0	970
Network size = 0	409,321	(47.1%)				
Network size between 1 and 5	208,494	(24.0%)	2.42	1.37	1	5
Network size between 6 and 20	142,883	(16.4%)	11.17	4.17	6	20
Network size between 21 and 50	67,920	(7.8%)	32.06	8.41	21	50
Network size above 50	40,207	(4.6%)	103.48	72.74	51	970
Network with a same technology field as patent	868,825		4.00	14.00	0	546
Network with different technology fields from patent	868,825		5.71	17.99	0	838
Number of patents developed by all patent inventors within last 5 years	868,825		6.40	17.09	0	542
Number of patent classes on patents developed by all patent inventors within last 5 years	868,825		2.04	2.64	0	44
Number of citations received (*output citations*)[1]	868,825		14.81	25.14	0	2193
Number of citations made (*input citations*)[2]	868,825		5.53	10.48	0	506
Mean age of input citations (measured in days)	714,627		2586.03	1347.86	0	9428
Age of the latest input citations (measured in days)	714,627		1430.62	1245.09	0	9428

[1,2] Self-citations are excluded.

Table 2: Patent Impact, Team Size and Network Size[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1)	(2)	(3)
Team size indicators			
Team size = 2 (*T2*)	0.063***	0.067***	0.065***
	(0.004)	(0.004)	(0.004)
Team size = 3 (*T3*)	0.124***	0.137***	0.133***
	(0.005)	(0.006)	(0.006)
Team size ≥ 4 (*T4*)	0.226***	0.240***	0.235***
	(0.007)	(0.008)	(0.008)
Network size			
Ln_network_size (*NS*)[2]	0.012***	0.020***	0.028***
	(0.004)	(0.004)	(0.004)
Interaction terms			
*T2 * NS*		-0.007**	-0.005*
		(0.003)	(0.003)
*T3 * NS*		-0.013***	-0.010**
		(0.004)	(0.004)
*T4 * NS*		-0.012**	-0.009*
		(0.005)	(0.005)
Other control variables			
Ln_5yr_patent_no[3]			-0.081***
			(0.010)
Ln_5yr_patent_class_no[4]			0.106***
			(0.013)
Patent class FE	423	423	423
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	358
R-squared	0.231	0.231	0.232
Obs.	868,825	868,825	868,825

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2]Ln_network_size is the logarithm of one plus the total number of inventors within two units of social distances for all team members, excluding ones from the same team. See Section 2 for details in variable construction.
[3]Ln_5yr_patent is the logarithm of the total number of patents developed by team members within last 5 years.
[4]Ln_5yr_patent_class_no is the logarithm of the total number of patent classes in patents developed by team members within last 5 years.

Table 3: Patent Impact Distribution, Team Size and Network Size[1]

Logit Regression DV	(1) Citation count in top 10% within the same year, patent class cohort	(2) Citation count = 0
Team size indicators		
Team size = 2 (*T2*)	0.124^{***}	-0.100^{***}
	(0.013)	(0.016)
Team size = 3 (*T3*)	0.276^{***}	-0.187^{***}
	(0.017)	(0.023)
Team size ≥ 4 (*T4*)	0.491^{***}	-0.310^{***}
	(0.020)	(0.026)
Network size		
Ln_network_size (*NS*)	0.065^{***}	-0.034^{***}
	(0.009)	(0.012)
Interaction terms		
*T2 * NS*	-0.016^{*}	0.003
	(0.010)	(0.011)
*T3 * NS*	-0.028^{***}	0.006
	(0.010)	(0.012)
*T4 * NS*	-0.023^{**}	-0.014
	(0.011)	(0.013)
Other control variables		
Ln_5yr_patent_no	-0.150^{***}	0.179^{***}
	(0.018)	(0.026)
Ln_5yr_patent_class_no	0.235^{***}	-0.203^{***}
	(0.025)	(0.034)
Patent class FE	423	423
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
Wald χ^2	151,678	44,295
Obs.	868,614	868,567

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Patents with patent class number 69 are dropped in the second column. This patent class is dropped because the patent class dummy does not converge after 30 rounds.

Table 4: Patent Impact, Team Size and Network Size (by Network Type)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1)	(2)	(3)
Team size indicators			
Team size = 2 (*T2*)	0.068^{***}	0.064^{***}	0.066^{***}
	(0.004)	(0.004)	(0.004)
Team size = 3 (*T3*)	0.138^{***}	0.131^{***}	0.135^{***}
	(0.005)	(0.006)	(0.006)
Team size ≥ 4 (*T4*)	0.246^{***}	0.234^{***}	0.239^{***}
	(0.007)	(0.008)	(0.008)
Network size			
Ln_network_size_same_class (*NS_SC*)[2]	0.022^{***}		0.014^{**}
	(0.005)		(0.006)
Ln_network_size_diff_class (*NS_DC*)[3]		0.030^{***}	0.023^{***}
		(0.005)	(0.005)
Interaction terms			
*T2 * NS_SC*	-0.011^{***}		-0.012^{***}
	(0.004)		(0.005)
*T3 * NS_SC*	-0.018^{***}		-0.018^{***}
	(0.006)		(0.006)
*T4 * NS_SC*	-0.017^{***}		-0.018^{***}
	(0.006)		(0.007)
*T2 * NS_DC*		-0.005	0.001
		(0.004)	(0.004)
*T3 * NS_DC*		-0.010^{**}	-0.001
		(0.005)	(0.005)
*T4 * NS_DC*		-0.009	-0.000
		(0.006)	(0.006)
Other control variables			
Ln_5yr_patent_no	-0.073^{***}	-0.074^{***}	-0.074^{***}
	(0.009)	(0.010)	(0.009)
Ln_5yr_patent_class_no	0.110^{***}	0.098^{***}	0.097^{***}
	(0.012)	(0.013)	(0.011)
Patent class FE	423	423	423
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	358
R-squared	0.231	0.232	0.232
Obs.	868,825	868,825	868,825

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Network sizes are separated into two: NS_SC and NS_DC.
[2]Ln_network_size_same_class is the logarithm of the size of the patent's affiliated network that shares the same technology class as the patent.
[3]Ln_network_size_diff_class is the logarithm of the size of the patent's affiliated network that has different patent classes from the patent.

Table 5: Patent Impact, Team Size and Network Size (by Knowledge Adoption Speed of Patent Classes)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Patent classes with lower than 50 percentile of citation age each year (*fast-evolving*)	(2) Patent classes with higher than 50 percentile of citation age each year (*slow-evolving*)
Team size indicators		
Team size = 2 (*T2*)	0.070***	0.055***
	(0.005)	(0.006)
Team size = 3 (*T3*)	0.145***	0.108***
	(0.008)	(0.008)
Team size ≥ 4 (*T4*)	0.257***	0.184***
	(0.010)	(0.011)
Network size		
Ln_network_size (*NS*)	0.029***	0.024***
	(0.004)	(0.005)
Interaction terms		
*T2 * NS*	-0.008**	-0.003
	(0.003)	(0.006)
*T3 * NS*	-0.017***	0.003
	(0.005)	(0.008)
*T4 * NS*	-0.018***	0.013
	(0.005)	(0.009)
Other control variables		
Ln_5yr_patent_no	-0.069***	-0.095***
	(0.010)	(0.015)
Ln_5yr_patent_class_no	0.106***	0.102***
	(0.015)	(0.017)
Patent class FE	363	351
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.244	0.159
Obs.	562,471	297,443

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Patent classes are separated into fast-evolving ones and slow-evolving ones according to the speed of knowledge adoption.

Table 6: Patent Impact, Team Size and Network Size (by Network Type and Knowledge Adoption Speed of Patent Classes)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Fast-evolving	(2) Slow-evolving
Team size indicators		
Team size = 2 (*T2*)	0.070***	0.056***
	(0.005)	(0.005)
Team size = 3 (*T3*)	0.144***	0.113***
	(0.007)	(0.008)
Team size ≥ 4 (*T4*)	0.257***	0.192***
	(0.009)	(0.011)
Network size		
Ln_network_size_same_class (*NS_SC*)	0.020***	0.002
	(0.007)	(0.007)
Ln_network_size_diff_class (*NS_DC*)	0.019***	0.029***
	(0.005)	(0.006)
Interaction terms		
*T2 * NS_SC*	-0.016***	-0.006
	(0.005)	(0.008)
*T3 * NS_SC*	-0.024***	-0.004
	(0.007)	(0.012)
*T4 * NS_SC*	-0.025***	-0.000
	(0.008)	(0.011)
*T2 * NS_DC*	0.003	-0.003
	(0.005)	(0.007)
*T3 * NS_DC*	-0.002	-0.000
	(0.006)	(0.007)
*T4 * NS_DC*	-0.004	0.008
	(0.007)	(0.009)
Other control variables		
Ln_5yr_patent_no	-0.065***	-0.087***
	(0.010)	(0.015)
Ln_5yr_patent_class_no	0.100***	0.092***
	(0.014)	(0.016)
Patent class FE	363	351
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.244	0.159
Obs.	562,471	297,443

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Patent Impact, Team Size and Network Size (with Firm FE)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) All	(2) Fast-evolving	(3) Slow-evolving
Team size indicators			
Team size = 2 (*T2*)	0.051***	0.055***	0.043***
	(0.004)	(0.006)	(0.006)
Team size = 3 (*T3*)	0.113***	0.116***	0.098***
	(0.006)	(0.007)	(0.009)
Team size ≥ 4 (*T4*)	0.209***	0.220***	0.168***
	(0.008)	(0.009)	(0.011)
Network size			
Ln_network_size_same_class (*NS_SC*)	0.022***	0.027***	0.013**
	(0.005)	(0.006)	(0.006)
Ln_network_size_diff_class (*NS_DC*)	0.021***	0.018***	0.023***
	(0.005)	(0.005)	(0.006)
Interaction terms			
*T2 * NS_SC*	-0.007*	-0.010**	-0.005
	(0.004)	(0.005)	(0.007)
*T3 * NS_SC*	-0.014***	-0.018***	-0.007
	(0.005)	(0.006)	(0.012)
*T4 * NS_SC*	-0.013**	-0.017**	-0.000
	(0.006)	(0.007)	(0.011)
*T2 * NS_DC*	0.003	0.006	-0.004
	(0.004)	(0.005)	(0.007)
*T3 * NS_DC*	0.003	0.005	0.002
	(0.005)	(0.006)	(0.008)
*T4 * NS_DC*	0.002	0.001	0.007
	(0.007)	(0.007)	(0.010)
Other control variables			
Ln_5yr_patent_no	-0.056***	-0.049***	-0.063***
	(0.007)	(0.008)	(0.012)
Ln_5yr_patent_class_no	0.077***	0.080***	0.070***
	(0.010)	(0.012)	(0.014)
Patent class FE	423	363	351
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	358
Firm FE	Yes	Yes	Yes
R-squared	0.373	0.388	0.354
Obs.	868,825	562,471	297,443

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Number of Citations Made, Team Size and Network Size[1]

OLS Regression DV: Ln(1 + number of input citations)	(1) Fast-evolving	(2) Slow-evolving
Team size indicators		
Team size = 2 (*T2*)	0.031***	0.038***
	(0.005)	(0.005)
Team size = 3 (*T3*)	0.073***	0.074***
	(0.007)	(0.007)
Team size ≥ 4 (*T4*)	0.138***	0.144***
	(0.010)	(0.012)
Network size		
Ln_network_size_same_class (*NS_SC*)	-0.009	0.005
	(0.006)	(0.009)
Ln_network_size_diff_class (*NS_DC*)	0.011**	0.026***
	(0.005)	(0.006)
Interaction terms		
*T2 * NS_SC*	-0.007	-0.008
	(0.004)	(0.009)
*T3 * NS_SC*	-0.014**	-0.009
	(0.006)	(0.010)
*T4 * NS_SC*	-0.012**	0.012
	(0.006)	(0.011)
*T2 * NS_DC*	-0.002	-0.009
	(0.004)	(0.007)
*T3 * NS_DC*	-0.002	-0.018**
	(0.005)	(0.007)
*T4 * NS_DC*	0.009	-0.020**
	(0.006)	(0.009)
Other control variables		
Ln_5yr_patent_no	-0.035***	-0.019
	(0.011)	(0.015)
Ln_5yr_patent_class_no	0.043***	0.010
	(0.016)	(0.019)
Patent class FE	363	351
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.274	0.279
Obs.	562,471	297,443

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Knowledge adoption Speed (Reciprocal of Age of the Latest Citation), Team Size and Network Size[1]

OLS Regression DV: Ln(reciprocal of (1 + age of the latest citation))	(1) Fast-evolving	(2) Slow-evolving
Team size indicators		
Team size = 2 (*T2*)	0.030***	0.030***
	(0.005)	(0.006)
Team size = 3 (*T3*)	0.059***	0.053***
	(0.007)	(0.007)
Team size ≥ 4 (*T4*)	0.104***	0.101***
	(0.008)	(0.009)
Network size		
Ln_network_size_same_class (*NS_SC*)	0.041***	0.028***
	(0.005)	(0.007)
Ln_network_size_diff_class (*NS_DC*)	0.049***	0.048***
	(0.004)	(0.005)
Interaction terms		
*T2 * NS_SC*	-0.004	-0.005
	(0.004)	(0.008)
*T3 * NS_SC*	-0.016***	-0.015*
	(0.005)	(0.008)
*T4 * NS_SC*	-0.017***	-0.012
	(0.006)	(0.008)
*T2 * NS_DC*	0.002	-0.005
	(0.004)	(0.007)
*T3 * NS_DC*	0.005	-0.007
	(0.004)	(0.007)
*T4 * NS_DC*	-0.001	-0.013*
	(0.005)	(0.007)
Other control variables		
Ln_5yr_patent_no	-0.102***	-0.063***
	(0.007)	(0.007)
Ln_5yr_patent_class_no	-0.003	-0.001
	(0.009)	(0.009)
Patent class FE	363	351
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.138	0.040
Obs.	475,073	233,766

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Different Measures of Patent Impact, Team Size and Network Size[1]

OLS Regression DV:	(1) Ln(1 + citation counts received by 2010, self-citations included)	(2) Ln(1 + citation counts received by 2010, excluding all citations from same CBSA)
Team size indicators		
Team size = 2 (*T2*)	0.067^{***}	0.062^{***}
	(0.004)	(0.004)
Team size = 3 (*T3*)	0.139^{***}	0.122^{***}
	(0.007)	(0.006)
Team size ≥ 4 (*T4*)	0.244^{***}	0.219^{***}
	(0.009)	(0.008)
Network size		
Ln_network_size (*NS*)	0.049^{***}	0.017^{***}
	(0.004)	(0.004)
Interaction terms		
*T2 * NS*	-0.007^{**}	-0.003
	(0.003)	(0.003)
*T3 * NS*	-0.014^{***}	-0.008^{**}
	(0.004)	(0.004)
*T4 * NS*	-0.009^{*}	-0.006
	(0.005)	(0.005)
Other control variables		
Ln_5yr_patent_no	-0.050^{***}	-0.093^{***}
	(0.010)	(0.008)
Ln_5yr_patent_class_no	0.080^{***}	0.121^{***}
	(0.013)	(0.011)
Patent class FE	423	423
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.216	0.226
Obs.	868,825	868,825

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 11: Patent Impact, Team Size and Network Size (in Broad Technology Field)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Fast-evolving	(2) Slow-evolving	(3) (1) – patents from "Computers and Communications"
Team size indicators			
Team size = 2 (*T2*)	0.072^{***}	0.057^{***}	0.064^{***}
	(0.005)	(0.006)	(0.006)
Team size = 3 (*T3*)	0.146^{***}	0.114^{***}	0.128^{***}
	(0.007)	(0.008)	(0.008)
Team size ≥ 4 (*T4*)	0.263^{***}	0.195^{***}	0.221^{***}
	(0.009)	(0.011)	(0.011)
Network size			
Ln_network_size_same_category (*NS_SCAT*)[2]	0.018^{***}	-0.009^{*}	0.013^{**}
	(0.005)	(0.005)	(0.006)
Ln_network_size_diff_category (*NS_DCAT*)[3]	0.030^{***}	0.064^{***}	0.041^{***}
	(0.007)	(0.008)	(0.008)
Interaction terms			
*T2 * NS_SCAT*	-0.012^{***}	0.001	-0.013^{**}
	(0.004)	(0.007)	(0.005)
*T3 * NS_SCAT*	-0.017^{***}	0.011	-0.014^{**}
	(0.006)	(0.009)	(0.006)
*T4 * NS_SCAT*	-0.025^{***}	0.013	-0.020^{**}
	(0.008)	(0.010)	(0.008)
*T2 * NS_DCAT*	0.003	-0.016^{**}	0.001
	(0.007)	(0.008)	(0.008)
*T3 * NS_DCAT*	-0.009	-0.028^{***}	-0.009
	(0.008)	(0.009)	(0.009)
*T4 * NS_DCAT*	-0.003	-0.017^{*}	-0.003
	(0.010)	(0.010)	(0.010)
Other control variables			
Ln_5yr_patent_no	-0.065^{***}	-0.084^{***}	-0.091^{***}
	(0.010)	(0.015)	(0.012)
Ln_5yr_patent_class_no	0.099^{***}	0.088^{***}	0.134^{***}
	(0.015)	(0.016)	(0.017)
Patent class FE	363	351	319
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	358
R-squared	0.245	0.159	0.230
Obs.	562,471	297,443	398,940

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2]Ln_network_size_same_category measures the network size for inventors sharing the same technology category as defined in Hall et al. (2001) with the patent in question.
[3]Ln_network_size_diff_category measures the network size for inventors with a different technology category with the patent.

Table 12: Patent Impact, Team Size and Network Size in Fast-evolving Patent Classes[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Excluding "Computers and Communications"	(2) Excluding "Electrics and electronics"	(3) (1) + (2)
Team size indicators			
Team size = 2 (*T2*)	0.063^{***}	0.065^{***}	0.053^{***}
	(0.006)	(0.006)	(0.008)
Team size = 3 (*T3*)	0.125^{***}	0.139^{***}	0.110^{***}
	(0.008)	(0.008)	(0.010)
Team size ≥ 4 (*T4*)	0.213^{***}	0.253^{***}	0.192^{***}
	(0.011)	(0.010)	(0.013)
Network size			
Ln_network_size_same_category (*NS_SCAT*)	0.012	0.021^{***}	0.008
	(0.008)	(0.007)	(0.009)
Ln_network_size_diff_category (*NS_DCAT*)	0.031^{***}	0.018^{***}	0.037^{***}
	(0.006)	(0.005)	(0.006)
Interaction terms			
*T2 * NS_SCAT*	-0.019^{***}	-0.013^{**}	-0.014^{*}
	(0.007)	(0.006)	(0.008)
*T3 * NS_SCAT*	-0.020^{***}	-0.028^{***}	-0.020^{**}
	(0.007)	(0.008)	(0.009)
*T4 * NS_SCAT*	-0.019^{**}	-0.022^{***}	-0.007
	(0.009)	(0.008)	(0.010)
*T2 * NS_DCAT*	0.003	0.003	0.001
	(0.006)	(0.005)	(0.007)
*T3 * NS_DCAT*	-0.000	-0.000	-0.001
	(0.007)	(0.006)	(0.008)
*T4 * NS_DCAT*	-0.001	-0.004	-0.002
	(0.008)	(0.007)	(0.010)
Other control variables			
Ln_5yr_patent_no	-0.090^{***}	-0.064^{***}	-0.098^{***}
	(0.012)	(0.011)	(0.014)
Ln_5yr_patent_class_no	0.131^{***}	0.093^{***}	0.130^{***}
	(0.016)	(0.016)	(0.019)
Patent class FE	319	315	271
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	358
R-squared	0.230	0.277	0.270
Obs.	398,940	439,482	275,951

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table 13: Patent Impact, Team Size and Network Size Using a Three-year Window Measure[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1)	(2)	(3)
Team size indicators			
Team size = 2 (*T2*)	0.071***	0.064***	0.066***
	(0.004)	(0.004)	(0.004)
Team size = 3 (*T3*)	0.143***	0.133***	0.137***
	(0.006)	(0.006)	(0.006)
Team size ≥ 4 (*T4*)	0.248***	0.234***	0.237***
	(0.008)	(0.009)	(0.009)
Network size			
Ln_network_size_same_class_3yr (*NS_SC_3yr*)[2]	0.038***		0.029***
	(0.005)		(0.005)
Ln_network_size_diff_class_3yr (*NS_DC_3yr*)[3]		0.035***	0.023***
		(0.004)	(0.004)
Interaction terms			
*T2 * NS_SC_3yr*	-0.014***		-0.017***
	(0.004)		(0.004)
*T3 * NS_SC_3yr*	-0.021***		-0.023***
	(0.005)		(0.006)
*T4 * NS_SC_3yr*	-0.018***		-0.021***
	(0.006)		(0.006)
*T2 * NS_DC_3yr*		-0.004	0.005
		(0.004)	(0.004)
*T3 * NS_DC_3yr*		-0.009**	0.002
		(0.005)	(0.005)
*T4 * NS_DC_3yr*		-0.007	0.003
		(0.006)	(0.006)
Other control variables			
Ln_5yr_patent_no	-0.092***	-0.084***	-0.093***
	(0.010)	(0.010)	(0.010)
Ln_5yr_patent_class_no	0.121***	0.097***	0.105***
	(0.013)	(0.013)	(0.012)
Patent class FE	423	423	423
Patent filing year FE	1980-1999	1980-1999	1980-1999
Patent CBSA FE	358	358	358
R-squared	0.223	0.223	0.223
Obs.	793,348	793,348	793,348

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The network measure in all regressions is constructed within a three-year time window as described in Section 2.
[2,3]Ln_network_size_same_class_3yr and ln_network_size_diff_class_3yr are similarly constructed as the original measures.

Table 14: Patent Impact, Team Size and Network Size Using a Three-year Window Measure (by Knowledge Adoption Speed of Patent Classes)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Fast-evolving	(2) Slow-evolving
<u>Team size indicators</u>		
Team size = 2 (*T2*)	0.071***	0.057***
	(0.006)	(0.006)
Team size = 3 (*T3*)	0.147***	0.117***
	(0.008)	(0.008)
Team size ≥ 4 (*T4*)	0.261***	0.184***
	(0.010)	(0.012)
<u>Network size</u>		
Ln_network_size_same_class_3yr (*NS_SC_3yr*)	0.034***	0.024***
	(0.006)	(0.006)
Ln_network_size_diff_class_3yr (*NS_DC_3yr*)	0.019***	0.027***
	(0.005)	(0.006)
<u>Interaction terms</u>		
*T2 * NS_SC_3yr*	-0.020***	-0.013*
	(0.005)	(0.007)
*T3 * NS_SC_3yr*	-0.027***	-0.017
	(0.006)	(0.011)
*T4 * NS_SC_3yr*	-0.025***	-0.013
	(0.007)	(0.011)
*T2 * NS_DC_3yr*	0.006	0.001
	(0.005)	(0.006)
*T3 * NS_DC_3yr*	0.002	0.003
	(0.006)	(0.007)
*T4 * NS_DC_3yr*	-0.003	0.016**
	(0.006)	(0.008)
<u>Other control variables</u>		
Ln_5yr_patent_no	-0.086***	-0.100***
	(0.011)	(0.015)
Ln_5yr_patent_class_no	0.109***	0.097***
	(0.015)	(0.017)
Patent class FE	363	348
Patent filing year FE	1980-1999	1980-1999
Patent CBSA FE	357	357
R-squared	0.234	0.144
Obs.	504,850	279,821

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix: Supplemental Tables

Table 151: High-impact Patent Indicator, Team Size and Network Size[1]

Logit Regression DV: Citation count in top 10% within the same year, patent class cohort	(1) Fast-evolving	(2) Slow-evolving
Team size indicators		
Team size = 2 (*T2*)	0.128^{***}	0.121^{***}
	(0.017)	(0.019)
Team size = 3 (*T3*)	0.284^{***}	0.264^{***}
	(0.022)	(0.025)
Team size ≥ 4 (*T4*)	0.523^{***}	0.439^{***}
	(0.025)	(0.031)
Network size		
Ln_network_size (*NS*)	0.060^{***}	0.084^{***}
	(0.011)	(0.015)
Interaction terms		
*T2 * NS*	-0.021^{*}	-0.010
	(0.011)	(0.016)
*T3 * NS*	-0.037^{***}	-0.013
	(0.012)	(0.019)
*T4 * NS*	-0.038^{***}	0.002
	(0.011)	(0.019)
Other control variables		
Ln_5yr_patent_no	-0.134^{***}	-0.188^{***}
	(0.020)	(0.025)
Ln_5yr_patent_class_no	0.239^{***}	0.237^{***}
	(0.030)	(0.033)
Patent class FE	363	351
Patent filing year FE	21	21
Patent CBSA FE	358	358
Wald χ^2	4,373	3,403
Obs.	562,156	297,185

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 162: Patent Impact, Team Size and Network Size in Fast-evolving Patent Classes[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Firms with fewer than 50 patents filed by 2000	(2) Firms with fewer than 100 patents filed by 2000	(3) Firms with more than 100 patents filed by 2000
Team size indicators			
Team size = 2 (*T2*)	0.062***	0.066***	0.070***
	(0.008)	(0.008)	(0.007)
Team size = 3 (*T3*)	0.147***	0.147***	0.133***
	(0.011)	(0.010)	(0.010)
Team size ≥ 4 (*T4*)	0.280***	0.274***	0.233***
	(0.013)	(0.013)	(0.012)
Network size			
Ln_network_size_same_class (*NS_SC*)	0.071***	0.058***	0.023***
	(0.011)	(0.010)	(0.007)
Ln_network_size_diff_class (*NS_DC*)	0.054***	0.047***	0.024***
	(0.011)	(0.009)	(0.007)
Interaction terms			
*T2 * NS_SC*	-0.025*	-0.025**	-0.015**
	(0.013)	(0.012)	(0.006)
*T3 * NS_SC*	-0.041***	-0.039***	-0.020***
	(0.015)	(0.013)	(0.007)
*T4 * NS_SC*	-0.088***	-0.076***	-0.014
	(0.015)	(0.013)	(0.009)
*T2 * NS_DC*	-0.019	-0.011	0.007
	(0.013)	(0.011)	(0.006)
*T3 * NS_DC*	-0.021	-0.010	0.003
	(0.014)	(0.012)	(0.007)
*T4 * NS_DC*	-0.023*	-0.011	0.001
	(0.014)	(0.012)	(0.008)
Other control variables			
Ln_5yr_patent_no	0.014	-0.012	-0.079***
	(0.013)	(0.013)	(0.011)
Ln_5yr_patent_class_no	0.026	0.055***	0.112***
	(0.017)	(0.017)	(0.016)
Patent class FE	362	362	361
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	350
R-squared	0.264	0.268	0.242
Obs.	162,201	196,214	366,257

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Samples are stratified by number of filed patents by firms.

Table 173: Patent Impact, Team Size and Network Size in Slow-evolving Patent Classes[1]

OLS Regression DV: Ln(1 + citation counts received by 2010)	(1) Firms with fewer than 50 patents filed by 2000	(2) Firms with fewer than 100 patents filed by 2000	(3) Firms with more than 100 patents filed by 2000
Team size indicators			
Team size = 2 (*T2*)	0.054^{***}	0.052^{***}	0.056^{***}
	(0.008)	(0.007)	(0.008)
Team size = 3 (*T3*)	0.105^{***}	0.104^{***}	0.115^{***}
	(0.011)	(0.011)	(0.011)
Team size ≥ 4 (*T4*)	0.177^{***}	0.175^{***}	0.199^{***}
	(0.015)	(0.014)	(0.015)
Network size			
Ln_network_size_same_class (*NS_SC*)	0.043^{***}	0.032^{**}	0.003
	(0.014)	(0.013)	(0.008)
Ln_network_size_diff_class (*NS_DC*)	0.053^{***}	0.046^{***}	0.035^{***}
	(0.012)	(0.010)	(0.007)
Interaction terms			
*T2 * NS_SC*	0.015	0.003	-0.008
	(0.019)	(0.017)	(0.008)
*T3 * NS_SC*	-0.009	0.011	-0.010
	(0.022)	(0.019)	(0.013)
*T4 * NS_SC*	-0.013	0.011	-0.005
	(0.023)	(0.020)	(0.012)
*T2 * NS_DC*	-0.012	-0.005	-0.001
	(0.015)	(0.013)	(0.007)
*T3 * NS_DC*	-0.002	0.009	-0.001
	(0.017)	(0.015)	(0.008)
*T4 * NS_DC*	0.010	0.012	0.007
	(0.017)	(0.015)	(0.010)
Other control variables			
Ln_5yr_patent_no	-0.004	-0.007	-0.106^{***}
	(0.013)	(0.012)	(0.017)
Ln_5yr_patent_class_no	0.013	0.015	0.105^{***}
	(0.017)	(0.016)	(0.019)
Patent class FE	351	351	348
Patent filing year FE	1980-2000	1980-2000	1980-2000
Patent CBSA FE	358	358	347
R-squared	0.161	0.163	0.167
Obs.	112,439	130,839	166,604

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Samples are stratified by number of filed patents by firms.

Table 18: Different Measure of Patent Impact, Team Size and Network Size (by Knowledge Adoption Speed of Patent Classes)[1]

OLS Regression DV: Ln(1 + citation counts received by 2010, excluding all citations from same CBSA)	(1) Fast-evolving	(2) Slow-evolving
Team size indicators		
Team size = 2 (*T2*)	0.065***	0.054***
	(0.006)	(0.006)
Team size = 3 (*T3*)	0.135***	0.096***
	(0.008)	(0.008)
Team size ≥ 4 (*T4*)	0.244***	0.163***
	(0.010)	(0.011)
Network size		
Ln_network_size (*NS*)	0.019***	0.013**
	(0.004)	(0.005)
Interaction terms		
*T2 * NS*	-0.005	-0.000
	(0.003)	(0.005)
*T3 * NS*	-0.014***	0.004
	(0.004)	(0.007)
*T4 * NS*	-0.015***	0.012
	(0.004)	(0.008)
Other control variables		
Ln_5yr_patent_no	-0.079***	-0.111***
	(0.009)	(0.013)
Ln_5yr_patent_class_no	0.118***	0.124***
	(0.013)	(0.015)
Patent class FE	363	351
Patent filing year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.238	0.156
Obs.	562,471	297,443

[1]Clustered standard errors in parentheses. Errors are clustered at firm level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 19: Knowledge Adoption Speed (Reciprocal of Mean Age of Citations), Team Size and Network Size[1]

OLS Regression DV: Ln(reciprocal of (1 + mean age of the citations)	(1) Fast-evolving	(2) Slow-evolving
Team size indicators		
Team size = 2 (*T2*)	0.012^{***}	0.001
	(0.003)	(0.003)
Team size = 3 (*T3*)	0.012^{***}	0.005
	(0.004)	(0.005)
Team size ≥ 4 (*T4*)	0.011^{**}	0.003
	(0.005)	(0.005)
Network size		
Ln_network_size_same_class (*NS_SC*)	0.037^{***}	0.015^{***}
	(0.003)	(0.004)
Ln_network_size_diff_class (*NS_DC*)	0.023^{***}	0.023^{***}
	(0.002)	(0.003)
Interaction terms		
*T2 * NS_SC*	-0.004	0.000
	(0.003)	(0.005)
*T3 * NS_SC*	-0.011^{***}	-0.006
	(0.003)	(0.005)
*T4 * NS_SC*	-0.013^{***}	-0.014^{***}
	(0.004)	(0.005)
*T2 * NS_DC*	0.001	0.000
	(0.003)	(0.004)
*T3 * NS_DC*	0.004	-0.001
	(0.003)	(0.004)
*T4 * NS_DC*	-0.002	0.006
	(0.003)	(0.004)
Other control variables		
Ln_5yr_patent_no	-0.036^{***}	-0.025^{***}
	(0.005)	(0.006)
Ln_5yr_patent_class_no	-0.013^{*}	0.003
	(0.007)	(0.007)
Patent Class FE	363	351
Patent application year FE	1980-2000	1980-2000
Patent CBSA FE	358	358
R-squared	0.173	0.243
Obs.	475,073	233,766

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 2: Non-Compete Clauses and R&D Inventor Team Size

Yanmin Yang

Department of Economics

Syracuse University

1. Introduction

The last decades have witnessed the rapid expansion of collaborations in the patent development process (Wutchy et al., 2007). While individual inventors developed more than 50% of patents in 1980, this ratio fell to around 30% by 2000. Meanwhile, large inventor teams contributed to an increasing share of patents (Yang, 2018). This phenomenon has been widely discussed in prior studies (Agrawal and Goldfarb, 2008; Jones, 2009; Ding et al., 2010; Singh and Fleming, 2010; Kim and Marschke, 2017).

Among these studies, it remains unclear whether changes in public policies or institutions might lead to an increase in inventor team size and what is the underlying mechanism if so. This paper provides an answer through focusing on a particular type of changes in law: the strengthening of non-compete clause enforceability in a state. Non-compete clauses are commonly used in high-tech firms to restrict the mobility of current employees towards rival firms. The stringency of the non-compete clause enforceability varies across states (Stuart and Sorensen, 2003). States with laws that are more stringent usually a lower job-switching rate (Fallick et al., 2006).

There are many channels through which the stringency of non-compete clauses enforceability can potentially affect inventor team sizes. The nature of non-compete clauses is to protect intellectual property, which incentivizes firms to increase R&D investment and expand inventor team sizes (Samilla and Sorensen, 2011; Kim and Marschke, 2017). However, the stringency in non-compete clause enforceability also has its disadvantages. It is obviously more difficult to hire researchers from other firms within the same state where non-compete clauses are commonly practiced. The use of non-compete clause could also reduce the benefit to learn from other firms since it impedes knowledge spillovers within the region (Fallick et al., 2006).

Both channels provide disincentives for R&D investment. Firms would cut down the investment if the cost of producing internal R&D outweighs its potential benefit.

This paper focuses in particular on the last channel. After the strengthening in non-compete clause enforceability, is it possible to compensate for the loss of knowledge researchers could otherwise access from the external network? The answer is yes. Yang (2018) found that inventor team size and network size are at least partial substitutes as sources of knowledge inputs in the patent development process. If this holds true, firms could increase internal R&D investment and expand inventor team size to substitute for the reduced external network size. This would compensate for the loss due to the reduced knowledge spillovers.

Using individual patent level data from the United States Patent and Trademark Office (USPTO) from 1975 to 1995, this paper examines the effect of the stringency of non-compete clause enforceability on two outcomes. The first outcome is patent inventor team size as measured by the number of patent inventors. The second outcome captures the ability for patent inventors to potentially access and acquire knowledge outside of their firms. It is measured by both network accessibility and external network size. The first outcome is measured by whether at least one patent inventor has recently changed jobs prior to patent development. The second one is the associated external network size for each patent. It is measured by the number of inventors within two units of distance outside of the patent's assignee firm following Yang (2018).

Through comparing the outcomes in Michigan versus states with similar non-compete clause enforceability, I identify the causal effects of Michigan Antitrust Reform Act (MARA) on both outcomes above using the difference-in-difference method. The enactment of MARA made non-compete clauses much easier to be implemented in Michigan and created an employerfriendly environment in this sense. The unexpected change on the non-compete clause enforceability due to MARA provides me the opportunity to study the causal relationship.

There are two main findings in my paper. First, MARA led to an increase in the average inventor team size. As compared with states with similar enforceability, patents are 4% less likely developed by individual inventors in Michigan from 1980 to 1990. Meanwhile, large teams with at least four inventors contribute 1.3% more to patent development in Michigan. This cannot be simply explained by the shifts in firm compositions. Old firms with at least one patent filed before 1980 experienced a sharp increase in inventor team sizes.

Second, the evidence of MARA on the network accessibility and the external network size is mixed. There is no clear decrease in external network sizes in Michigan after MARA. However, firms are less likely to hire inventors from other firms and incorporate them to develop new patents. From 1980 to 1990, the likelihood of having one patent developed by an inventor who has recently changed jobs raised from 12.9% to 16.0% in non-Michigan states after MARA. However, changes in Michigan are much smaller and the percentage only increased by 0.5%. Put it differently, the patent development process in Michigan was less likely to access information and knowledge from other firms within the same state.

Through exploring the causal relationship, this paper contributes to the literature on non-compete clauses and optimal rule of team formation in the following ways. First, my study demonstrates that increasing the stringency of non-compete clause enforceability can affect R&D investment through inventor team size expansions. Prior studies on non-compete clauses mainly focused on their effects on wage, job mobility and job training (Fallick et al., 2006; Marx et al., 2009; Starr et al. 2017). There is only a small but growing literature studying the effects of non-compete clauses on entrepreneurship and R&D investment (Samila and Sorensen, 2013; Conti, 2014; Carlino, 2017; Kim and Marschke, 2017). This paper provides supporting evidence that the strengthening of non-compete clause enforceability would provide incentives for firms to expand inventor team sizes. In this way, firms compensate for the loss of knowledge or information learned from the shrinking external network.

Second, I also shed light on the optimal rules of team formation for entrepreneurs. The literature on team formation are mainly studied under the framework of Becker and Murphy (1992): the optimal team size is the trade-off between the decreasing specialization benefits versus the increasing communication costs. This paper extends this idea and provides a new perspective builds on Yang (2018). Since inventor teams and networks are viewed as substitutes, this paper examines whether an exogenous interference in network formation would be compensated by the increase in inventor team sizes. Instead of focusing on specialization benefits, my approach accounts for the benefits from the professional network. This perspective will give us different views on how R&D teams could be optimally designed.

As far as the author knows, Kim and Marschke (2017) is the closest paper to my research. While my empirical findings are consistent with theirs, the underlying logic to explain the phenomenon are largely different. They implicitly assumed the stringency of non-compete clause enforceability would increase inventor team size due to the nature of intellectual property protection. Therefore, firms have more incentives to build large teams since it is less likely to leak information from current employees under the new circumstance. However, if firms frequently hire inventors from other firms to develop new patents, they may find it difficult to adjust inventor team size due to the increasing hiring cost. My paper finds that old firms, especially those relied heavily on hiring outside inventors for patent development experienced a higher increase in patent inventor team size after MARA. This is better explained by Yang

(2018): these firms have greater incentives to increase inventor team size to compensate for the great loss of knowledge inflow due to the shrinking external network.

The rest of the paper is outlined as follows. Section 2 provides the background information for MARA in 1985. Data and variables are introduced in Section 3. Section 4 introduces the empirical model and testable hypotheses. Main results are reported and discussed in Section 5. Section 6 concludes the paper.

2. Michigan Antitrust Reform Act (MARA) in 1985

Michigan has a long history against the use of non-compete clauses. Statute 445.761 in Act 329 of 1905 greatly constrained the use of non-compete clauses:

"All agreements and contracts by which any person, co-partnership or corporation promises or agrees not to engage in any avocation, employment, pursuit, trade, profession or business, whether reasonable or unreasonable, partial or general, limited or unlimited, are hereby declared to be against public policy and illegal and void." (as cited in Carlino, 2017)

This act remained in use until the Michigan Antitrust Reform Act (MARA) became effective in March 1985.[15] The main purpose of MARA was to increase competition and reduce the abuse of monopoly power in the market. However, the repeal of the original statute allowed employers to enforce non-compete clauses without violating laws. Marx et al. (2009) conducted a thorough study and found no evidence that non-compete clause legislation served as the main motivation of the act. It was also not a response to firm's lobbying behavior. In fact, non-compete clauses have not been mentioned in the law until December 1987: if a non-compete clause in the agreement is "found to be unreasonable in any respect, a court may limit the agreement to render it reasonable" (as cited in Carlino, 2017). As a result, MARA is used as a

[15] The Michigan Antitrust Reform Act is also known as Act 274 of 1984.

quasi-natural experiment to study changes in non-compete clauses enforceability (Marx et al., 2009; Carlino, 2017).

MARA also provided a substantive change in the enforceability of non-compete clauses. Changes in non-compete clause enforceability are uncommon and most changes happened while non-compete clauses have been well established in the laws (Marx et al., 2009; Garmaise, 2011). Unlike other states, "Michigan is the only state we know to have clearly and inadvertently reversed its enforcement policy in the past century" (Marx et al., 2009). This provides an extra advantage of using MARA over using changes of non-compete clause enforceability in other states. With such a reverse in the enforcement policy, MARA is expected to have a large effect on inventor team size and external network size.

3. Data and variables

The data used in this paper is the same as in Yang (2018), which includes all US patents from the United States Patent and Trade Office (USPTO) from 1975 to 2010. It provides detailed information on each patent: name and residential address of each inventor, patent class, patent filing year, citations made and received by the patent, etc. I also complemented it with firm level identifier from NBER patent project.[16]

The sample period in this paper ranges from 1976 to 2000. In the main analysis, I restricted the data set to include only utility patents in Michigan and other states in the contiguous US with similar non-compete clause enforceability during MARA.[17] This choice of

[16] This data can be downloaded from https://sites.google.com/site/patentdataproject/Home. If there are multiple assignees, the first one is used.
[17] According to Malsberger(1996), the following states are included: Alaska, California, Connecticut, Minnesota, Montana, North Dakota, Nevada, Oklahoma, Washington, and West Virginia. Alaska is not included in the data set since I only focus states in the contiguous US.

control groups is similar to Marx et al. (2009). I will discuss different ways to select control groups in Section 4.

There are two sets of main outcome measures. The first one is team size. For each patent, team size is defined as the total number of inventors. It is further grouped into four categories: one inventor, two inventors, three inventors, and four inventors and above. The second outcome measure includes two network measures: external network size and network accessibility. External network size is similar to the network size constructed in Yang (2018). For each patent, it counts the total number of inventors who satisfies the following conditions. First, these inventors should come from a different firm from patent inventors. Second, they have to be close to inventors in the patent inventor team. More specifically, these inventors should be within two units of distance away from any inventors in the team. Third, qualifying inventors should have worked closely with any inventors in the team around or prior to patent development. In particular, inventors in the external network should have worked with patent inventors or their collaborators within two years of the patent's filing year. If the external network size is larger, inventors are more likely to have interactions and knowledge access from other firms.[18]

The second measure, network accessibility, is an indicator that equals one if a patent includes at least one inventor who was recently hired from another firm within the same state. This measure is an extension of within-state job mobility of inventors used by Marx et al. (2009). It captures how likely a patent, instead of an inventor, incorporates knowledge and information from other firms within the same state.

To construct this measure, I first identified inventors who changed jobs following Marx et al. (2009). For each inventor, all patents developed by him/her are listed in a chronological

[18] Refer to Yang (2018) for details on variable construction.

order. An inventor is considered to have changed a job if the patent he developed does not belong to the same firm as the previous one. As is illustrated in Figure 1, inventor A developed patents 1 to 5 from 1982 to 1988. Starting in 1985, patents 3 to 5 are assigned to company K instead of company J. Since I cannot identify when exactly inventor A moved from company J to K., I used the filing year of the newly developed patent (patent 3) as the year inventor A changed jobs.

The network accessibility is constructed based on the move measure. Network accessibility is set to one for all patents filed with such an *external* inventor in the year he/she moved. Therefore, patents 3 and 4 in Figure 1 has network accessibility equal to one. The network accessibility for the last patent developed in 1988 is zero unless another patent inventor has just changed a job. Since this paper focuses on non-compete clause which mainly restricts labor mobility within the same state, the network accessibility measure is set to one only if inventors move within a state. I will address the concern of between-state movements of inventors in Section 5.3.

The following variables are used as control variables. I calculated the number of inventors in a CBSA in the same technology field to control for the local growth of inventors. Similar to Yang (2018), for each patent, I also calculated the number of patents developed last year by any inventors in the team, and the number of patent classes listed on these patents. They are used to capture differences in patenting histories. Summary statistics for all variables are reported in levels in Table 1.

4. Empirical methods

4.1 Difference-in-difference model and testable hypotheses

The main empirical strategy used in the analysis is difference-in-difference (DID). Considering patent i developed in CBSA in year t with patent class c:

$$\text{outcome}_{i,c,CBSA,t} = \alpha\ 1(t \geq 1986) * 1(\text{CBSA in MI}) + \theta_c + \gamma_{CBSA} + \eta_t + \beta X_{i,c,CBSA,t} + \varepsilon_{i,c,CBSA,t} \quad (1)$$

There are two outcome measures as described in Section 3. One measures the inventor team size for the patent, and the other captures the ability for the patent inventors to incorporate knowledge outside of the inventor firms. On the right-hand side of equation (1), the interaction term includes both a time dummy and a CBSA dummy. The time dummy describes the MARA reform. It equals one only if the patent is filed after 1986.[19] The CBSA dummy is set to one only if most of the patent inventors lived in CBSAs within the Michigan state.

To reduce the concern of omitted variables, I included patent class fixed effect θ_c, filing year fixed effect η_t and CBSA fixed effect γ_{CBSA} to address heterogeneities across technology fields, years and geographic areas. $X_{i,c,CBSA,t}$ represents all control variables described in Section 3.

α is the parameter of interest in this study. It captures the differences between outcomes in Michigan before and after 1986, comparing with a selected control group of other states. For example, if the patent inventor team size grew faster (slower) after MARA in Michigan, α is expected to be positive (negative).

The sign of α for each outcome can be predicted by Yang (2018). Yang (2018) found that inventor team size and network size are partial substitutes in determining patent impact. In

[19] Although MARA took place in March 1985, year 1986 was used as the time period dummy in this paper following Marx et al. (2009).

particular, the marginal contribution of team size is higher when the associated network size is smaller. Since MARA unexpectedly made non-compete clauses much easier to be implemented in Michigan after 1986, it created barriers for inventor job mobility and network formation across firms. According to Yang (2018), this shrink in external network size can be compensated through expansion of inventor team size. The prediction of the sign of α is summarized in the following hypotheses:

Hypothesis 1: *Team size will grow in Michigan after MARA in 1986, compared with states with similar non-compete clause enforceability. Therefore, α is expected to be positive if the outcome variable is inventor team size.*

Hypothesis 2: *Network accessibility and network size will drop after MARA in 1986, compared with states with similar enforceability. Hence, α is expected to be negative if the outcome variable is network accessibility or network size.*

One of the key issues in using difference-in-difference method is the selection of control groups. Prior studies used different control groups: ten other states with similar non-compete clause enforceability (Marx et al., 2009), California and North Dakota (Carlino, 2017; Kim and Marschke, 2017), all other states in the US (Carlino, 2017). In the main analysis, I will follow Marx et al. (2009) and choose states with similar non-compete clause enforceability as the control groups.[20] These states will be labeled as NCC states. Results with other control groups will be reported in Section 5.3.

4.2 Synthetic control method

The ambiguity in the selection of control units in the difference-in-difference method has been widely discussed in recent years. To address the issue, Abadie and Gardeazabal (2003), and

[20] These ten states are referred as "NCC states" in the following paragraphs.

Abadie et al. (2010) proposed a statistical solution: the synthetic control method. This method constructs the counterfactual using a weighted sum of outcomes over all possible control units, where the weight is selected based on the match between the observation and the counterfactual *before* treatment. Compared to the traditional difference-in-difference method, the synthetic control method addresses the concern of correlation in unobserved error terms with a small sample size and limited post-treatment period (Doudchenko and Imbens, 2016). It has also been widely used in comparative case studies where the number of treated units is small and the key variable of interest is measured at an aggregate level (Abadie and Gardeazabal, 2003; Abadie et al., 2010; Liu, 2015; Acemoglu et al., 2016).

For several reasons, the synthetic control method cannot be directly applied in the main analysis. First, it requires strongly balanced panel data (Abadie et al., 2010). However, the current data set does not allow me to construct a panel at firm level. Firms often have missing information in several years because no patents are filed in these years. Second, it is almost computationally infeasible to apply the synthetic control method with a large number of control units. Finally, there is also an aggregation problem. While the patent class is well defined for each individual patent, it is difficult to define the patent class for the average inventor team size at firm level.

Despite these problems, I am going provide a simply case study using the synthetic control method in Section 5.1. I will illustrate changes in average inventor team size in Detroit-Warren-Dearborn and a suitable synthetic counterfactual one. Detroit is chosen since it accounts for more than 50% of patents developed in Michigan from 1976 to 2000. As a comparison, I will also demonstrate changes in mean team size in Cleveland-Elyria.[21]

5. Results

5.1 Case study: Detroit versus Cleveland

This section illustrates two case studies using the synthetic control method. Figures 3(a) and 3(b) illustrate the effect of MARA on average inventor team sizes in Detroit-Warren-Dearborn and Cleveland-Elyria, respectively. The vertical axis measures the average team size and the horizontal axis describes the filing year of patents. A dashed vertical line is used to illustrate MARA reform.

In both figures, the counterfactual lines match quite well with the observations before 1986. However, there is a large divergence after 1986 in Figure 2(a). Average inventor team size in Detroit-Warren-Dearborn is much larger during that period. The gap between the counterfactual average team size and the real one is around 0.1 on average after 1986. It is equivalent to 6% in team size in 1986 when the average team size was approximately 1.8.

Although Figure 2(a) presents an increase in average team size in Detroit-Warren-Dearborn after MARA, this pattern is not observed in nearby cities. In Figure 2(b), the difference between the average team size in Cleveland-Elyria and the counterfactual one is small after MARA. The observed average team size is slightly higher from 1984 to 1988 in the Cleveland area, but the gap is small in scale and does not happen during MARA period.

---

[21] In both exercises, I include land area, number of inventors within the CBSA and the average team size every other two years from 1976 to 1984 as observed covariates.

5.2 Empirical results

5.2.1 The effect of MARA on inventor team size

5.2.1.1 Basic results

This section discusses empirical results using the difference-in-difference method. Table 2 reports the results of MARA on inventor team sizes using an ordered logit model. It is used since the dependent variable is an ordinal dummies representing inventor team size. All regressions include patent class fixed effects, patent filing year fixed effects and CBSA fixed effects. Columns 1 to 5 present empirical results with different time windows.

The coefficients the interaction terms are positive and significantly different from zero across all columns. This suggests that MARA increases the inventor team size in Michigan, compared with other states with similar non-compete clause enforceability. To better understand these results, I calculate the average marginal effect of MARA on inventor team size without filing year fixed effects and CBSA fixed effects in the period of 1980 to 1990.[22] Instead, I included the reform time dummy and a state dummy representing the Michigan state. This implicitly sets the base level to be the average inventor team size in the other states before MARA. Under the new setting, the probability that patents are developed by an individual inventor drops by 9% in non-Michigan states after MARA. In Michigan, however, the gap is even larger: patents are 13% less likely to be developed by a single inventor. Meanwhile, large teams with at least four inventors are 2.3% more likely to develop patents in the control group while this percentage rises to 3.6% in Michigan after MARA.

Observing positive values of the interaction terms does not guarantee the increase in inventor team size is driven by MARA. The nature of the difference-in-difference estimator is to

---

[22] The marginal effect of MARA on inventor team sizes is not estimable when year fixed effects and CBSA fixed effects are included. The MARA reform indicator is perfect collinear with filing year dummies.

compare the average treatment effect pre and post treatment in the treated and control groups. More specifically, when the sample period ranges from 1976 to 1994, it is less clear whether the observed positive coefficient is driven by changes around 1986.

To address the issue, I replaced the MARA reform dummy with year dummies for each year in equation (1) and conducted empirical analyses using the new model. This is a generalized difference-in-difference method. For each year, the interaction term measures the relative difference between average inventor team size in Michigan and other states to that in the base level (year 1976).

The coefficients of interaction terms are illustrated in Figure 3. It is straightforward that the coefficient turns positive around 1985 and keeps growing afterward. Furthermore, it is worth noting that most coefficients before 1984 are small and not significantly different from zero. This is consistent with the parallel trend assumption required for difference-in-difference method. It is not drive by the specific functional form: this pattern holds when estimated with a linear model.

The increase in inventor team size in Michigan might also be driven by a common trend in areas around Michigan. To address this concern, I followed Marx et al. (2009) and conducted a similar exercise for Ohio as a placebo test. If there are common trends in states nearby, inventor team sizes in Ohio are also expected to increase. Corresponding results using ordered logit regressions are illustrated in Figure 4. Although there are increases in coefficients in the late 1990s, there are no clear changes in the late 1980s. The coefficients of interaction terms in Ohio shortly after 1986 are insignificant and negative in sign. All of these empirical results reinforce my findings in Section 5.1: the growth in average inventor team size in Michigan is very likely driven by MARA.

5.2.1.2 Inventor team sizes in old firms

There might be concerns that the observed results are driven by shifts in the composition of firms. Small inventor teams might disproportionately come from small firms and startups. If the strengthening of non-compete clause enforceability impedes the entry of startups, previous findings might be explained by the reduction of small firms in Michigan. The share of individual inventors falls simply because there are fewer small startups.

This concern is addressed in the following way. First, there is no clear evidence that changes in non-compete clause enforceability would affect new firm entries and startups in the literature (Starr et al., 2015). Second, I can rule out the effect from new and small firms by focusing on a subsample of relatively *old* firms that existed after 1986. In this analysis, old firms are defined as those with at least one patent filed both before 1980 and after 1986. Around 60% of patents from the original sample remains after restricting the sample to include only patents from old firms.

Results for these old firms are reported in Table 3. To further address the issue of heterogeneities in firms, I include firm fixed effect in all regressions, along with standard controls and fixed effects used in previous models.[23] The coefficients of interaction terms are positive across all columns. Columns 1 to 2 suggest that MARA has a small but positive effect on inventor team size. In a longer time span, the coefficients are more significant and much larger (columns 3 to 5). This is suggestive that MARA increases the average inventor team size for old firms in Michigan. Therefore, my findings in the previous section cannot be simply explained by shifts in firm compositions.

[23] It is infeasible to include in my basic results since there are too many firms in total.

I also conducted the empirical analyses on old firms using the generalized difference-in-difference method by replacing the MARA reform dummy with filing year dummies in equation (1). Figures 5 and 6 illustrate the coefficients of the interaction terms similar to those in Figures 4 and 5. It is clear that the coefficients turn positive and significant around 1986 in Michigan, and such a pattern is not observed in Ohio. Although there are great increases in the coefficients after 1995 in Ohio, it would not be a concern for my previous empirical analyses. All of my previous results are estimated using samples ranging from 1976 to 1994, hence the difference-in-difference estimator could not be driven by changes after 1995.

5.2.1.3 Heterogeneous effects by old firm's reliance on 'external' inventors

In this section, I am going to examine how changes in non-compete clause enforceability would affect inventor team sizes by focusing on two types of old firms. I separated old firms into two types based on their reliance on external inventors[24] for patent development. For each old firm, I calculated the share of patents with network accessibility that equals one from 1976 to 1985. Firms with higher shares of patents developed by external inventors than state average are considered to have *high-reliance* on these inventors.[25] In other words, these firms rely heavily on hiring inventors from other firms for new patent development. In contrast, *low-reliance* firms refer to those with lower shares of patents developed by external inventors.

Tables 4 and 5 report regression results for low-reliance and high-reliance firms, respectively. MARA has little immediate effects on inventor team sizes for low-reliance firms. In the first three columns of Table 4, the coefficients of interaction terms are close to 0 and

[24] External inventors refer to inventors who have recently changed a job and moved to the firm in question.
[25] Firms with fewer than 10 patents filed during that period of time are dropped.

highly insignificant. As I expand the sample period, these coefficients increase monotonically and become significant in column 5 of Table 4.

The results for high-reliance firms have a different pattern. MARA has significant positive effects in the very short run (columns 1 to 3) while its long run effects are smaller and insignificant. This finding can be better explained by Yang (2018). Changes in non-compete clause enforceability create great barriers for external inventor network formations in a sudden. New channels of knowledge adoption are required immediately to compensate for the loss of knowledge transferred through inventor networks. Therefore, firms with higher-reliance on external inventors would have greater incentive to expand team sizes.

To summarize, I find consistent evidence that inventor team size grows in Michigan after MARA was enacted in 1985, compared with states with similar non-compete clause enforceability. It is not driven by shifts in firm compositions and it cannot be simply explained by common regional trends. Moreover, firms with higher reliance on external inventors for patent development present a higher and more significant growth in inventor team sizes, at least in the short run.

5.2.2 The effect of MARA on network accessibility and external network size

Section 5.2.1 presents that the enactment of MARA increases inventor team sizes. According to Yang (2018), this phenomenon can be explained because inventor team sizes and inventor network sizes are partial substitutes in the patent development process. Since MARA allows employer to enforce non-compete clauses at a lower cost, it would impede inventor job mobility and external network formations within Michigan. This direct channel is going to be checked in this section.

Table 6 reports the results of MARA reform on the network accessibility measure for all firms. The coefficients of the interaction terms are negative across all columns and become significantly negative within a six-year window (1982-1988). It suggests that network accessibility in Michigan drops quickly after the enactment of MARA. The effect on network accessibility also depreciates over time. The absolute values of the interaction terms in columns 4 and 5, where a longer time span is specified, are smaller than those in columns 2 and 3.

I also calculated the marginal effect of MARA on the network accessibility for the sample period ranging from 1980 to 1990.[26] The likelihood of having one patent developed by an external inventor increases from 12.9% to 16.0% in non-Michigan states after MARA. However, the growth rate in Michigan is only 0.5%. In other words, Michigan has not experienced as fast an increase in network accessibility as other states with similar non-compete clause enforceability. This result also holds true for old firms alone (Table 7).

These findings are similar to Marx et al. (2009). The enactment of MARA created great barriers for inventors to switch jobs within Michigan. Therefore, it also impeded inventor network formations through changing jobs. Although results in the first columns of Tables 6 and 7 do not follow such a pattern, they could result from lags between MARA announcement and contract renewals.

Table 8 presents the effect of MARA on the external network size for all firms. The external network size is positively correlated with the network accessibility, and the correlation coefficient between them is around 0.4. All coefficients of the interaction terms in Table 8 are

[26] To calculate the marginal effect, I dropped the filing year dummies and CBSA dummies and re-estimated the model. The coefficient of the interaction term in the new model is -0.236, which is not largely different from the estimates in the current model.

insignificant and small in magnitude. There is also no clear relationship between the enactment of MARA and external network sizes for old firms (Table 9).

There are two potential explanations for this finding. First, Yang (2018) found that network size and inventor team size are strong substitutes mainly in fast-evolving industries. The major industry in Michigan, especially Detroit is automobile. It might not be considered as a fast-evolving industry and the reduction in external network size is less clear. Second, it might also be explained by the selection in hiring inventors. Due to the increasing hiring costs, managers and entrepreneurs might be more likely to hire more talented inventors from other firms. If the professional network size from these talented inventors is larger, the average network size associated with each patent might not be affected. As a summary, my empirical results confirm that MARA reduced the likelihood of hiring inventors from other firms for patent development. However, there is no clear evidence that MARA reduces the associated external network size.

5.3 Robustness

Two robustness exercises are conducted in this section. First, I re-estimated equation (1) with different sets of control groups. To simplify the analysis, I only report results with the longest time span (1976 – 1994). This provides me with enough observations under all circumstances.

Tables 10 and 11 report the results of MARA on inventor team size and network accessibility with different control groups, respectively. Column 1 presents the original result where states with similar non-compete clause enforceability are treated as control groups. In column 2, I compare inventor teams in Michigan with those in California, where it is well known that non-compete clauses are almost impossible to implement. In column 3, I exclude California from the ten selected states used in the main analysis. This helps me to check whether my main findings are mainly driven by observations from California: concerns might arise since California has many small startups with small research teams. In the last column, I include all states in the contiguous US.

The coefficients of interaction terms have expected positive signs across all columns. The largest coefficient comes from column 2 when the control groups are inventor teams from California. This raises concerns for studies that only consider observations from California as control groups. The increase in inventor team size could be driven due to the uniqueness in entrepreneurship structure in California. Results of the enactment of MARA on network accessibility are also robust with different control groups (Table 11).

Second, there might be concerns about the definition of network accessibility. The network accessibility in the main analysis does not account for out-of-state movements. This channel is important since Marx et al. (2015) observed that many inventors left Michigan after MARA. To address the concern, I constructed a generalized network accessibility measure: if a patent is developed by at least one inventor who has recently changed jobs, the generalized measure is set to one.

Tables 12 and 13 report the effect of MARA on the generalized network accessibility for all firms and old firms only, respectively. All results are similar to those in Table 6 and 7. Results are still robust if I exclude patents developed with at least one inventor who changed jobs between states (not reported).

6. Conclusions

Using the quasi-natural experiment of MARA, this paper studies the causal effect of changes in non-compete clause enforceability on inventor team size and inventor network accessibility and external network size. The mean inventor team size in Michigan grows after the implementation of the new policy, as compared to states with similar non-compete clause enforceability. In a ten-year window from 1980 to 1990, I find that Michigan is 1.3% more likely to have a patent developed by large teams with at least four inventors. Meanwhile, it is 5% less likely to have patents created through individual inventors. This result holds for old firms, especially those who rely heavily on external inventors for patent development.

Findings on the effect of MARA on external network size and network accessibility are mixed. After the enactment of MARA, patents in Michigan are less likely developed by external inventors who have changed jobs in recent years. While the likelihood that one patent has at least one external inventor increases by 3.1% in non-Michigan states after 1986, it is only increased by 0.5% in Michigan. Therefore, inventors in Michigan is less likely to access knowledge and information from other firms during patent development process. However, there are no clear findings about the relationship between MARA and the external network size measure as defined by Yang (2018). Empirical results are small in magnitude and highly insignificant. This might come from measurement errors or selections in hiring decisions.

This paper helps us to rethink how non-compete clauses would affect R&D investments. The nature of non-compete clauses is to protect intellectual property, which would provide incentives for firms to invest in innovations. However, non-compete clauses would also reduce knowledge spillovers from job movers through impeding the job mobility across firms. This provides a disincentive for R&D investment as the reduction in external knowledge transfers reduces the marginal benefit of R&D investment. Based on Yang (2018), this paper points out a possibility to compensate the loss due to non-compete clauses through expanding inventor team sizes. This provides supporting evidence for non-compete clauses to incentivize R&D investments.

This paper has also pointed a new channel through which patent inventor team size could potentially grow. Large inventor teams have been playing an increasingly important role in patent development. Undoubtedly, the great decrease in communication costs, the emergence of the Internet, etc. have fostered a broader level of collaborations. In this paper, I find that the strengthening of non-compete clause enforceability could also increase the inventor team size. This finding can be explained by Yang (2018), which discovered that inventor teams and networks serve as substitutes in the patent development process. Therefore, the reduction in network accessibility due to changes in non-compete clause enforceability can be partially compensated by increasing inventor team size. Understanding this trade-off would shed light on the optimal allocations and team formations within such a "networked world", which can be helpful for entrepreneurs and managers.

For future research, it would be important to understand and take account of the hiring decisions from firms. The strengthening in non-compete clauses enforceability would not only change entrepreneurs' R&D investment incentives but also affect their hiring decisions. Under the new circumstance, it will be more expensive to hire inventors from other rivalry firms. This could potentially affect how and whom managers and entrepreneurs would like to employ. Accounting for the hiring decision could help us to better understand the micro-foundations of inventor team adjustments.

References

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review, 93*(1), 113-132.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association, 105*(490), 493-505.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., & Mitton, T. (2016). The value of connections in turbulent times: Evidence from the United States. *Journal of Financial Economics*, *121*(2), 368-391.

Agrawal, A., Goldfarb, A., & Teodoridis, F. (2016). Understanding the changing structure of scientific inquiry. *American Economic Journal: Applied Economics, 8*(1), 100-128.

Becker, G. S., & Murphy, K. M. (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics, 107*(4), 1137-1160.

Carlino, G. A. (2017). Do Non-Compete Covenants Influence State Startup Activity? Evidence from the Michigan Experiment.

Conti, R. (2014). Do non-competition agreements lead firms to pursue risky R&D projects?. *Strategic Management Journal, 35*(8), 1230-1248.

Doudchenko, N., & Imbens, G. W. (2016). *Balancing, regression, difference-in-differences and synthetic control methods: A synthesis* (No. w22791). National Bureau of Economic Research.

Fallick, B., Fleischman, C. A., & Rebitzer, J. B. (2006). Job-hopping in Silicon Valley: some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics, 88*(3), 472-481.

Garmaise, M. J. (2011). Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *The Journal of Law, Economics, and Organization*, *27*(2), 376-425.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?. *The Review of Economic Studies, 76*(1), 283-317.

Kim, J., & Marschke, G. (2017). Teams in R&D: Evidence from US Inventor Data. Working paper.

Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, *87*, 25-41.

Marx, M., Strumsky, D., & Fleming, L. (2009). Mobility, skills, and the Michigan non-compete experiment. *Management Science, 55*(6), 875-889.

Marx, M., Singh, J., & Fleming, L. (2015). Regional disadvantage? Employee non-compete agreements and brain drain. *Research Policy, 44*(2), 394-404.

Samila, S., & Sorenson, O. (2011). Noncompete covenants: Incentives to innovate or impediments to growth. *Management Science, 57*(3), 425-438.

Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality?. *Management Science, 56*(1), 41-56.

Starr, E., Balasubramanian, N., & Sakakibara, M. (2017). Screening spinouts? How noncompete enforceability affects the creation, growth, and survival of new firms. *Management Science, 64*(2), 552-572.

Starr, E. (2017). Consider this: Training, wages, and the enforceability of covenants not to compete.

Stuart, T. E., & Sorenson, O. (2003). Liquidity events and the geographic distribution of entrepreneurial activity. *Administrative Science Quarterly*, *48*(2), 175-201.

Yang, Y. (2018). Innovation in a Networked World: Inventor Team, Social Network and Patent Impact. Working paper

Figure 1: Within State Inventor Movement

Patent	Inventor	Application Date	Company	Accessibility
001	A	Sep 1982	J	0
002	A	Jan 1984	J	0
003	A	Mar 1985	K	1
004	A	Oct 1985	K	1
005	A	Feb 1988	K	0

Figure 2(a): Impact of MARA on Average Team Size in Detroit-Warren-Dearborn

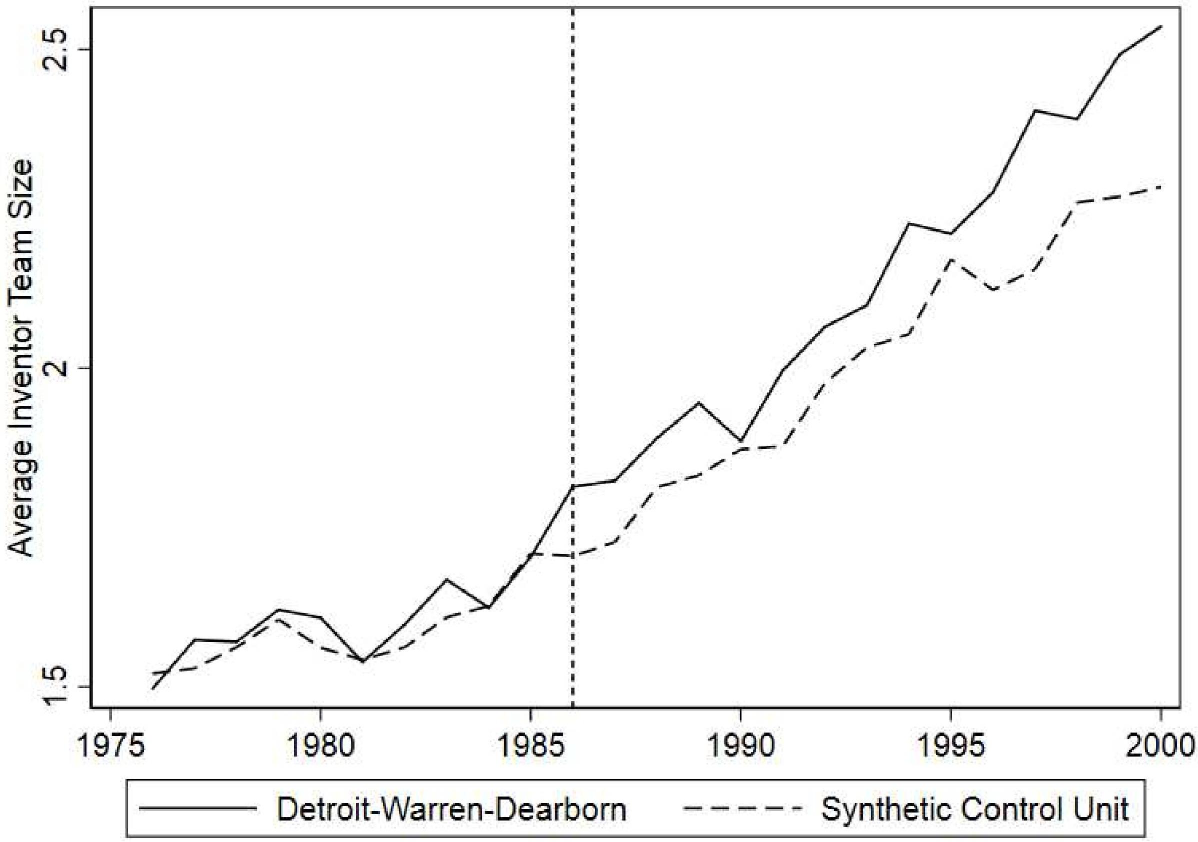


Figure 2(b): Impact of MARA on Average Team Size in Cleveland-Elyria

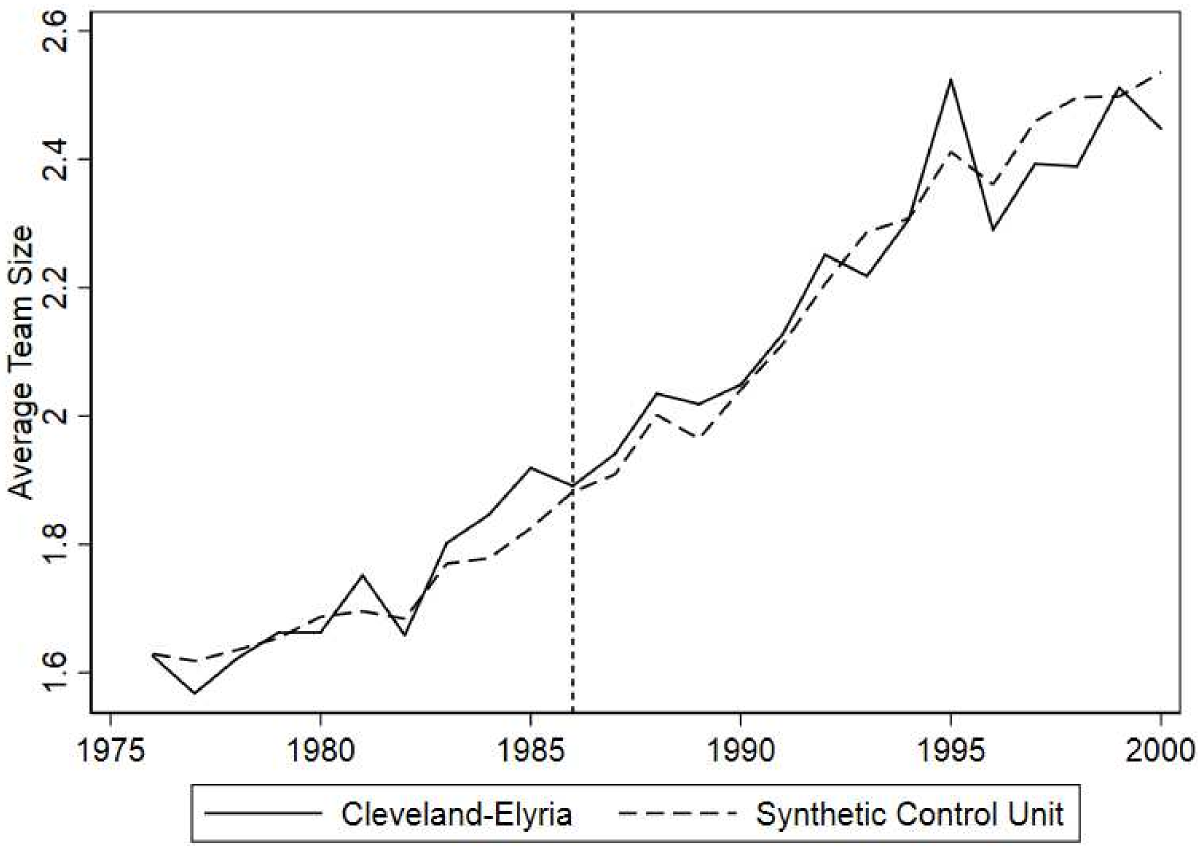
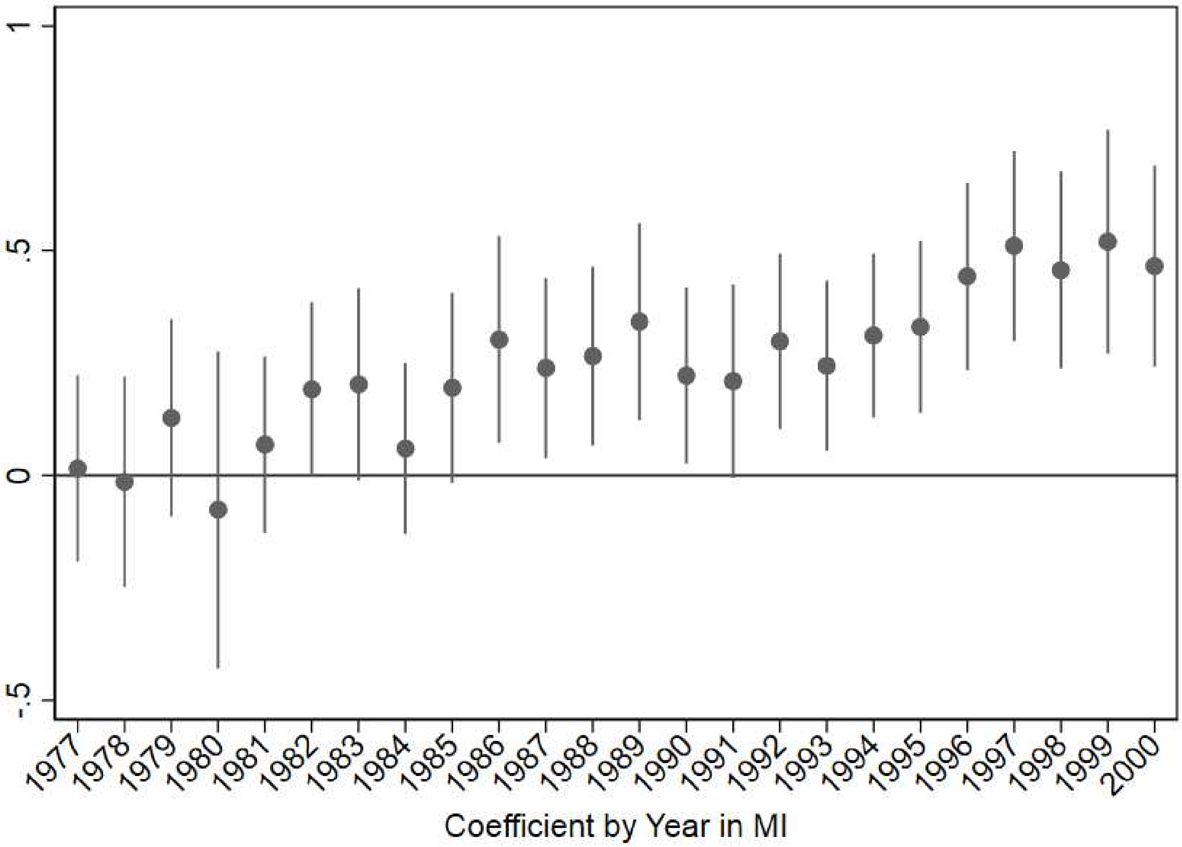
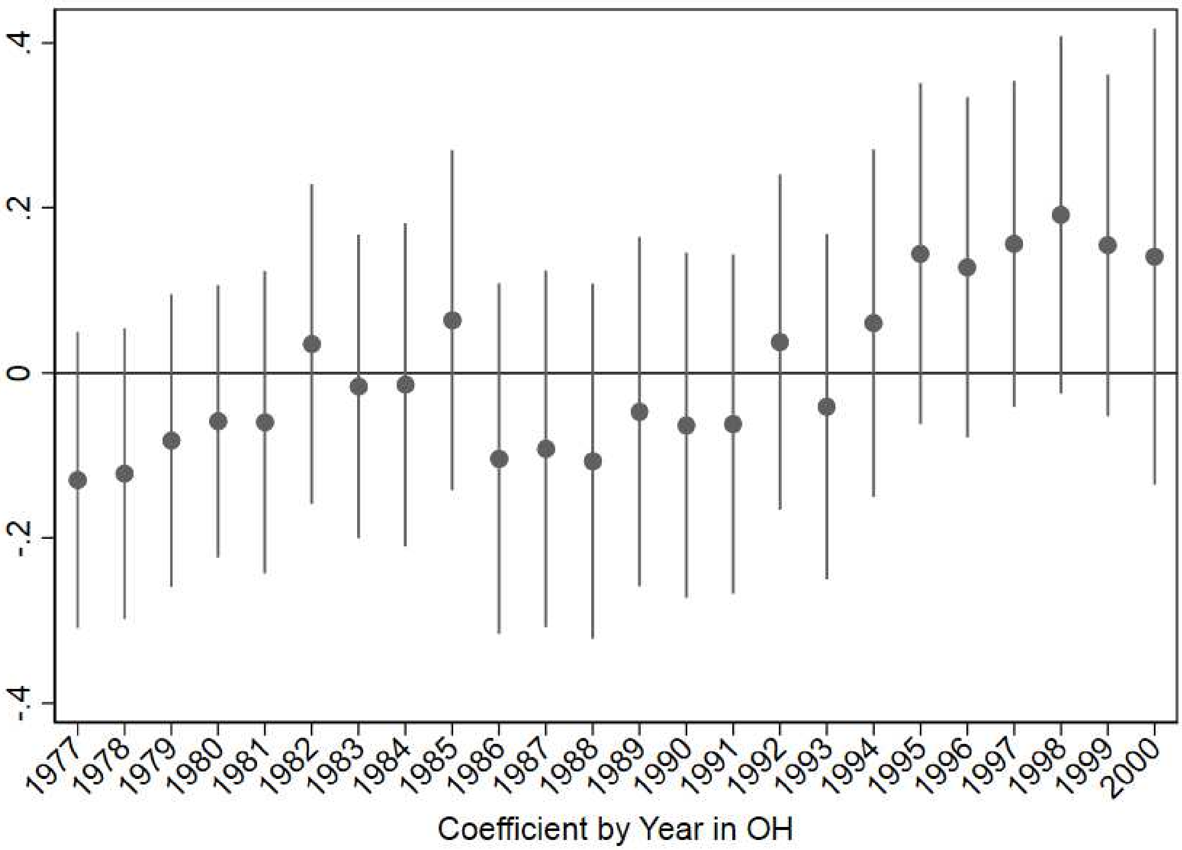


Figure 3: Coefficients of Interaction Terms for Michigan versus NCC States (all firms)[1]



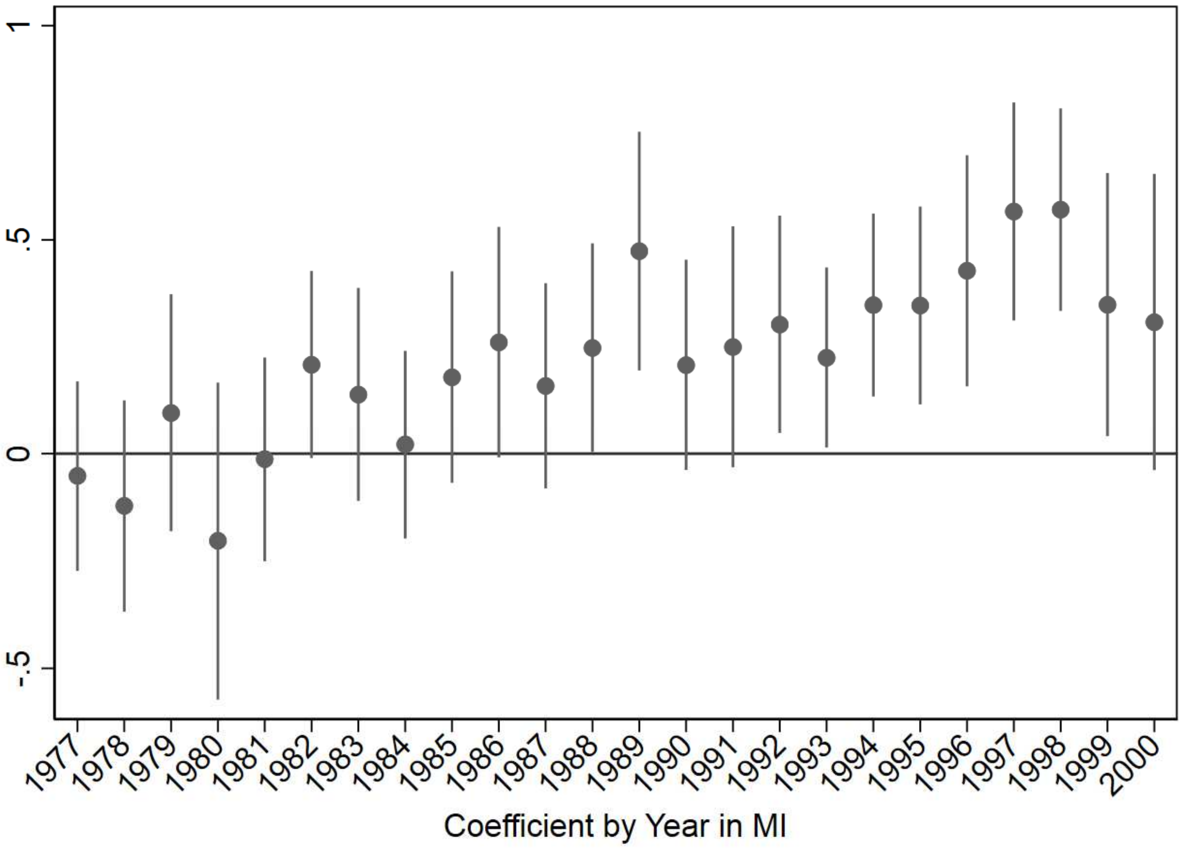
[1]The reported coefficient is α in equation (1). Regression includes patent class FE, application year FE, CBSA FE, logarithm of number of inventors in broad technology category in CBSA, patenting experience controls. Errors are clustered at firm level.

Figure 4: Coefficients of Interaction Terms for Ohio versus NCC States (all firms)1



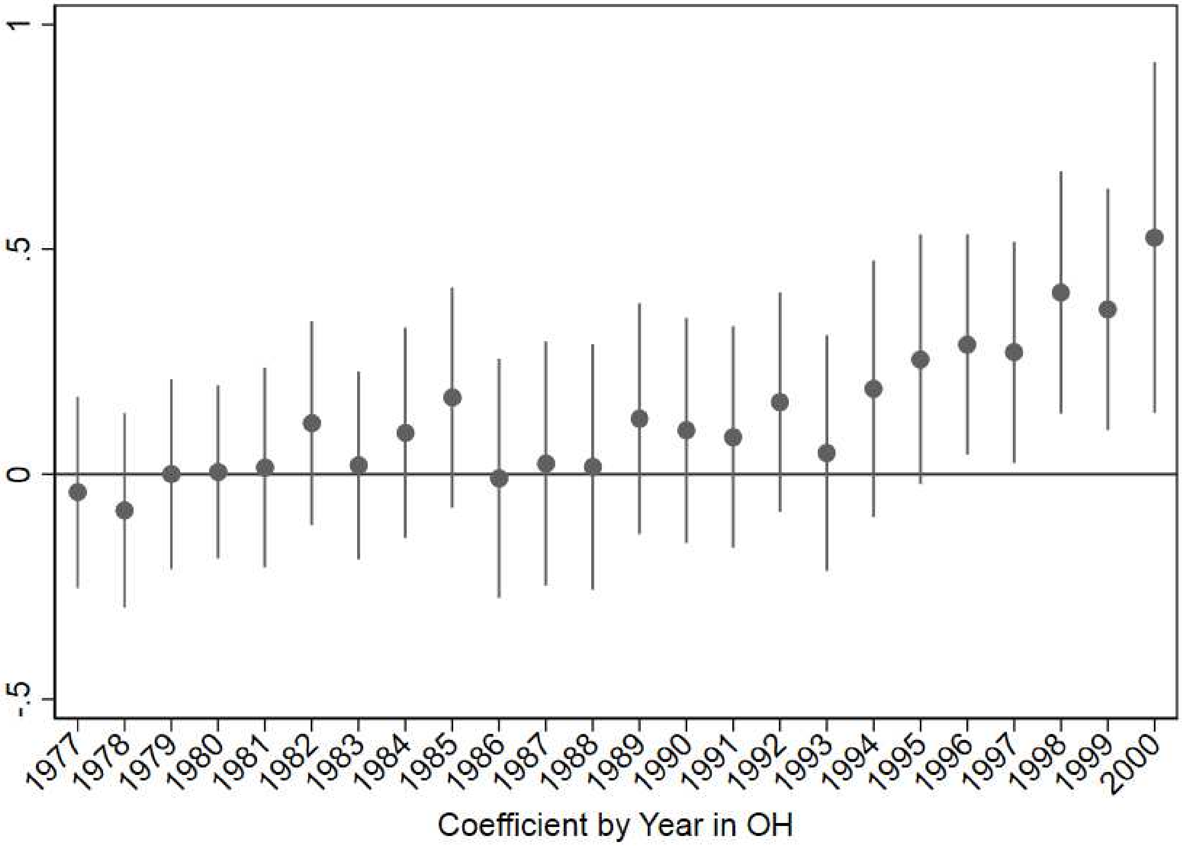
[1]The reported coefficient is α in equation (1). Regression includes patent class FE, application year FE, CBSA FE, logarithm of number of inventors in broad technology category in CBSA, patenting experience controls. Errors are clustered at firm level.

Figure 5: Coefficients of Interaction Terms for Michigan versus NCC States (old firms)[1]



[1]The reported coefficient is α in equation (1). Samples only include firms with at least one patent before 1980 and after 1986. Regression includes patent class FE, application year FE, CBSA FE, logarithm of number of inventors in broad technology category in CBSA, patenting experience controls. Errors are clustered at firm level.

Figure 6: Coefficients of Interaction Terms for Ohio versus NCC States (old firms)1



[1]The reported coefficient is α in equation (1). Samples only include firms with at least one patent before 1980 and after 1986. Regression includes patent class FE, application year FE, CBSA FE, logarithm of number of inventors in broad technology category in CBSA, patenting experience controls. Errors are clustered at firm level.

Table 1: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Accessibility	306,847	0.19	0.39	0	1
Team size	306,847	2.20	1.47	1	34
1 inventor	120,919 (39.41%)				
2 inventors	91,246 (29.80%)				
3 inventors	50,393 (16.42%)				
4 inventors and above	44,109 (14.37%)	4.99	1.57	4	34
Network size from other firms	306,847	1.92	7.70	0	260
1(Year >= 1986)	306,847	0.78	0.41	0	1
1(CBSA in Michigan)	306,847	0.15	0.36	0	1
Number of inventors in technology field in CBSA	306,847	678.71	986.87	1	5029
Number of patents developed last year	306,847	1.89	6.02	0	219
Number of patent classes involved last year	306,847	0.90	1.41	0	24

Table 2: Team Size in Michigan versus NCC States (all firms)[1]

Ordered logit model DV: team size indicators	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	0.177**	0.098*	0.150**	0.168**	0.202***
	(0.090)	(0.058)	(0.063)	(0.069)	(0.066)
Ln_inventor_CBSA_cat[2]	0.179***	0.185***	0.169***	0.137***	0.113***
	(0.040)	(0.029)	(0.024)	(0.022)	(0.020)
Ln_patent_1yr[3]	0.195**	0.201***	0.041	0.071	0.109
	(0.097)	(0.068)	(0.099)	(0.091)	(0.076)
Ln_patent_class_1yr[4]	0.390***	0.379***	0.554***	0.540***	0.511***
	(0.121)	(0.084)	(0.112)	(0.100)	(0.086)
Patent Class FE	395	408	414	416	418
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	78	81	83	84	85
Obs.	21,511	51,634	84,815	121,385	163,831

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2]Ln_inventor_CBSA_cat is the logarithm of the number of inventors in the same CBSA and broad technology field as the patent inventors in the same filing year.
[3]Ln_patent_1yr is the total number of patents developed by patent inventors in the team last year.
[4]Ln_patent_class_1yr is the total number of patent classes listed on patents developed by patent inventors nn the team last year.

Table 3: Team size in Michigan versus NCC States (old firms)[1]

Ordered logit model DV: team size indicators	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	0.130	0.075	0.165*	0.191*	0.239**
	(0.134)	(0.074)	(0.092)	(0.105)	(0.098)
Ln_inventor_CBSA_cat	0.244***	0.188***	0.161***	0.125***	0.112***
	(0.055)	(0.038)	(0.032)	(0.033)	(0.033)
Ln_patent_1yr	0.328***	0.313***	0.166	0.149	0.204*
	(0.107)	(0.074)	(0.146)	(0.132)	(0.113)
Ln_patent_class_1yr	0.373***	0.342***	0.466***	0.494***	0.436***
	(0.139)	(0.090)	(0.157)	(0.139)	(0.118)
Patent Class FE	377	394	399	401	404
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	66	72	76	76	79
Firm FE	1078	1343	1521	1634	1682
Obs.	14,428	33,670	54,410	75,957	98,825

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Team Size in Michigan versus NCC States (old firms with low-reliance on external inventors)[1]

Ordered logit model DV: team size indicators	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	-0.101	0.002	0.096	0.168	0.239*
	(0.184)	(0.100)	(0.123)	(0.142)	(0.127)
Ln_inventor_CBSA_cat	0.148*	0.117**	0.103***	0.069*	0.073*
	(0.076)	(0.048)	(0.037)	(0.041)	(0.044)
Ln_patent_1yr	0.109	0.094	-0.022	-0.048	0.021
	(0.121)	(0.089)	(0.164)	(0.150)	(0.135)
Ln_patent_class_1yr	0.559***	0.544***	0.640***	0.676***	0.603***
	(0.156)	(0.111)	(0.184)	(0.167)	(0.152)
Patent Class FE	345	370	382	388	389
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	56	59	59	60	62
Firm FE	254	268	270	271	271
Obs.	7,378	17,444	28,515	40,439	52,831

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Team Size in Michigan versus NCC States (old firms with high-reliance on external inventors)[1]

Ordered logit model DV: team size indicators	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	0.397*	0.243*	0.269**	0.162	0.186
	(0.204)	(0.131)	(0.122)	(0.123)	(0.116)
Ln_inventor_CBSA_cat	0.316***	0.256***	0.223***	0.197***	0.144***
	(0.088)	(0.052)	(0.041)	(0.036)	(0.038)
Ln_patent_1yr	0.501***	0.545***	0.526***	0.489***	0.507***
	(0.186)	(0.094)	(0.081)	(0.069)	(0.052)
Ln_patent_class_1yr	0.333	0.176	0.165*	0.217**	0.196***
	(0.281)	(0.125)	(0.099)	(0.085)	(0.071)
Patent Class FE	323	354	367	376	382
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	43	48	54	56	58
Firm FE	257	262	263	264	264
Obs.	5,705	13,023	20,351	27,545	35,438

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Network Accessibility in Michigan versus NCC states (all firms)[1]

Logit model DV: network accessibility	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	-0.062	-0.377^{***}	-0.281^{***}	-0.211^{**}	-0.233^{**}
	(0.157)	(0.122)	(0.105)	(0.100)	(0.104)
Ln_inventor_CBSA_cat	-0.102	-0.081^{*}	-0.045	-0.050	-0.021
	(0.063)	(0.047)	(0.038)	(0.033)	(0.031)
Ln_patent_1yr	-0.479^{**}	-0.364^{**}	-0.475^{***}	-0.575^{***}	-0.502^{***}
	(0.226)	(0.149)	(0.125)	(0.111)	(0.102)
Ln_patent_class_1yr	0.923^{***}	0.782^{***}	0.881^{***}	0.976^{***}	0.902^{***}
	(0.254)	(0.167)	(0.140)	(0.125)	(0.116)
Patent Class FE	395	408	414	416	418
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	78	81	83	84	85
Obs.	20,895	51,354	84,532	121,080	163,581

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Network Accessibility in Michigan versus NCC States (old firms)[1]

Logit model DV: network accessibility	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	-0.358^{*}	-0.613^{***}	-0.368^{**}	-0.210	-0.188
	(0.189)	(0.178)	(0.159)	(0.152)	(0.165)
Ln_inventor_CBSA_cat	-0.117	-0.061	-0.026	-0.017	-0.012
	(0.098)	(0.079)	(0.058)	(0.052)	(0.049)
Ln_patent_1yr	-0.460^{*}	-0.272^{*}	-0.361^{**}	-0.469^{***}	-0.458^{***}
	(0.251)	(0.163)	(0.144)	(0.133)	(0.131)
Ln_patent_class_1yr	0.946^{***}	0.706^{***}	0.785^{***}	0.902^{***}	0.901^{***}
	(0.286)	(0.176)	(0.155)	(0.144)	(0.148)
Patent Class FE	377	394	399	401	404
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	66	72	76	76	79
Obs.	12,987	32,271	53,342	74,770	97,873

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: External Network Size in Michigan versus NCC States (all firms)[1]

Linear model DV:external network size	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	0.039	0.001	-0.006	-0.014	-0.017
	(0.049)	(0.040)	(0.031)	(0.026)	(0.026)
Ln_inventor_CBSA_cat	0.031^{***}	0.005	0.009	0.015^{**}	0.026^{***}
	(0.011)	(0.009)	(0.007)	(0.007)	(0.007)
Ln_patent_1yr	0.090^{**}	0.108^{***}	0.073^{**}	0.067^{*}	0.115^{***}
	(0.043)	(0.038)	(0.037)	(0.035)	(0.036)
Ln_patent_class_1yr	0.075	0.062	0.106^{***}	0.124^{***}	0.086^{**}
	(0.047)	(0.039)	(0.037)	(0.035)	(0.036)
Patent Class FE	395	408	414	416	418
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	78	81	83	84	85
Firm FE	4847	8923	12785	16715	20958
Obs.	21,511	51,634	84,815	121,385	163,831

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: External Network Size in Michigan versus NCC States (old firms)[1]

Linear model DV:external network size	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	0.052	0.014	0.006	-0.005	-0.009
	(0.053)	(0.044)	(0.033)	(0.028)	(0.028)
Ln_inventor_CBSA_cat	0.035^{***}	0.007	0.013^{*}	0.019^{**}	0.026^{***}
	(0.011)	(0.010)	(0.008)	(0.008)	(0.008)
Ln_patent_1yr	0.061	0.107^{***}	0.079^{**}	0.076^{**}	0.103^{***}
	(0.043)	(0.040)	(0.038)	(0.038)	(0.039)
Ln_patent_class_1yr	0.084^{*}	0.042	0.078^{**}	0.088^{**}	0.066^{*}
	(0.047)	(0.041)	(0.038)	(0.038)	(0.038)
Patent Class FE	377	394	399	401	404
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	66	72	76	76	79
Firm FE	1078	1343	1521	1634	1682
Obs.	14,428	33,670	54,410	75,957	98,825

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Team size in Michigan versus Different Control Groups (all firms during 1976 - 1994)[1]

Ordered logit model DV: team size indicators	(1) NCC states	(2) CA only	(3) (1) - CA	(4) All other states in the contiguous US
(Year >= 1986)* (CBSA in MI)	0.202***	0.221***	0.141**	0.199***
	(0.066)	(0.063)	(0.070)	(0.075)
Ln_inventor_CBSA_cat	0.113***	0.095***	0.142***	0.139***
	(0.020)	(0.022)	(0.029)	(0.013)
Ln_patent_1yr	0.109	0.088	0.074	0.250***
	(0.076)	(0.078)	(0.125)	(0.045)
Ln_patent_class_1yr	0.511***	0.491***	0.583***	0.380***
	(0.086)	(0.088)	(0.135)	(0.049)
Patent Class FE	418	413	416	422
Patent CBSA FE	85	41	59	358
Obs.	163,831	112,167	81,595	574,263

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Network Accessibility in Michigan versus Different Control Groups (all firms during 1976 - 1994)[1]

Logit model DV: network accessibility	(1) NCC states	(2) CA only	(3) (1) - CA	(4) All other states in the contiguous US
(Year >= 1986)* (CBSA in MI)	-0.233**	-0.197*	-0.284**	-0.321***
	(0.104)	(0.107)	(0.115)	(0.102)
Ln_inventor_CBSA_cat	-0.021	0.002	-0.104**	-0.050**
	(0.031)	(0.032)	(0.042)	(0.021)
Ln_patent_1yr	-0.502***	-0.492***	-0.393***	-0.420***
	(0.102)	(0.118)	(0.119)	(0.058)
Ln_patent_class_1yr	0.902***	0.895***	0.825***	0.792***
	(0.116)	(0.137)	(0.131)	(0.063)
Patent Class FE	418	413	416	422
Patent CBSA FE	85	41	59	358
Obs.	163,581	111,816	81,215	573,852

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Generalized Network Accessibility in Michigan versus NCC States (all firms)[1]

Logit model DV: generalized network accessibility	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	-0.089	-0.283***	-0.208**	-0.204**	-0.229**
	(0.139)	(0.104)	(0.093)	(0.089)	(0.094)
Ln_inventor_CBSA_cat	-0.094	-0.092**	-0.062*	-0.047	-0.022
	(0.061)	(0.043)	(0.035)	(0.031)	(0.028)
Ln_patent_1yr	-0.393*	-0.268*	-0.399***	-0.490***	-0.403***
	(0.214)	(0.146)	(0.121)	(0.107)	(0.099)
Ln_patent_class_1yr	0.829***	0.679***	0.808***	0.902***	0.825***
	(0.239)	(0.163)	(0.136)	(0.119)	(0.112)
Patent Class FE	395	408	414	416	418
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	78	81	83	84	85
Obs.	21,130	51,440	84,641	121,200	163,602

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Generalized Network Accessibility in Michigan versus NCC States (old firms)[1]

Logit model DV: generalized network accessibility	(1)	(2)	(3)	(4)	(5)
(Year >= 1986)* (CBSA in MI)	-0.408**	-0.426***	-0.290**	-0.222*	-0.228
	(0.167)	(0.143)	(0.136)	(0.129)	(0.142)
Ln_inventor_CBSA_cat	-0.117	-0.089	-0.078	-0.046	-0.036
	(0.090)	(0.067)	(0.053)	(0.046)	(0.042)
Ln_patent_1yr	-0.442**	-0.184	-0.338**	-0.422***	-0.377***
	(0.219)	(0.154)	(0.142)	(0.128)	(0.124)
Ln_patent_class_1yr	0.910***	0.585***	0.747***	0.836***	0.819***
	(0.248)	(0.172)	(0.158)	(0.141)	(0.142)
Patent Class FE	377	394	399	401	404
Patent application year FE	1984-1986	1982-1988	1980-1990	1978-1992	1976-1994
Patent CBSA FE	66	72	76	76	79
Obs.	13,529	32,912	53,972	75,539	98,432

[1]Clustered standard errors in parentheses. Errors are clustered at firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

VITA

AUTHOR:	Yanmin Yang
EMAIL:	yyang109@syr.edu
WEBSITE:	https://yyang109.expressions.syr.edu/

EDUCATION:

2014	M.S., Fudan University, Shanghai, China
2011	B.S., Fudan University, Shanghai, China

RESEARCH EXPERIENCE AND EMPLOYMENT:

2014 – 2018	Teaching Assistant, Department of Economics, Syracuse University
2015 – 2016	Research Assistant for Professor Abdulaziz Shifa

AWARDS AND HONORS:

2014 – 2019	Syracuse University Graduate Assistantship, Syracuse University