Collaborative Knowledge Creation: Evidence from Japanese Patent Data

Tomoya Mori[*,†] Shosei Sakaguchi[‡,§]

December 1, 2018

Abstract

This paper presents micro-econometric evidence for the mechanism of collaborative knowledge creation at the level of individual researchers based on Berliant and Fujita's (2008) model. The key driver for developing new ideas is found to be the exchange of differentiated knowledge among collaborators. To stay creative, inventors seek opportunities to shift their technological expertise to unexplored niches by utilizing the differentiated knowledge of new collaborators, in addition to their own stock of knowledge. While collaborators' differentiated knowledge raises all the average cited count, average technological novelty and the quantity of patents for which an inventor contributes to the development, it has the largest impact on the average novelty among the three.

Keywords: Knowledge creation, Collaboration, Differentiated knowledge, Technological novelty, Technological shift, Recombination, Patents, Network, Strategic interactions

JEL Classification: D83, D85, O31, R11, C33, C36

Acknowledgements: Earlier versions of this paper were presented at Jinan University, Kyoto University, RIETI, the University of Tokyo, the 13th meeting of Urban Economic Association.The authors are particularly indebted to Marcus Berliant, Jorge De la Roca, Gilles Duranton, Masahisa Fujita, Tadao Hoshino, Daishoku Kanahara, Shin Kanaya, Kentaro Nakajima, Yoshihiko Nishiyama, Ryo Okui and Akihisa Shibata for their constructive comments. This research was conducted as part of the project "An empirical framework for studying spatial patterns and causal relationships of economic agglomeration" undertaken at RIETI, and it received partial financial support from JSPS Grants-in-Aid for Scientific Research Grant Numbers 15H03344, 16K13360 and 17H00987.

[*]Institute of Economic Research, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan. E-mail: mori@kier.kyoto-u.ac.jp.

[†]Research Institute of Economy, Trade and Industry (RIETI), 11th floor, Annex, Ministry of Economy, Trade and Industry (METI) 1-3-1, Kasumigaseki Chiyoda-ku, Tokyo 100-8901, Japan.

[‡]Research Fellow, Japan Society for the Promotion of Science.

[§]Graduate School of Economics, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: sakaguchi@e.u-tokyo.ac.jp.

1 Introduction

Knowledge creation has been a key factor in various aspects of economic modeling. Some of the new ideas result in innovations, which fuel economic growth.[1] The structure of market and competition may be subject to the extent of diffusion and imitations of invented technologies.[2] Furthermore, the concentration of research and development (R&D) activities is the defining feature of the largest cities.[3] Yet, the mechanism of knowledge creation – at the ultimate micro level of individual inventors – has not been explicitly specified in these strands of the literature. Empirical studies are necessarily scarce.[4] In this study, we investigate data on Japanese patents applied between 1995 and 2009. Given that about 90% of patents were developed in collaborations, we focus on causalities in collaborative knowledge creation based primarily on the microeconomic model proposed by Berliant and Fujita (2008).[5]

We consider two complementary measures of output from a given patent project: one reflecting *quality* based on forward citation counts and the other reflecting *novelty* based on the timing of each patent application in the relevant technological category. The productivity of a given inventor is then defined in terms of the quality/novelty-adjusted count of patents in which this inventor participated.

Under either measure, our data indicate the presence of substantial downward pressure on inventor productivity: fewer than half of inventors with above-median productivity in a given period maintain at least the same relative productivity in the next period, where some top inventors stay highly productive, while some inferior inventors overthrow superior ones and climb the productivity ladder.[6] Overall, the substantial churning of relative productivities of inventors is observed over time.

The extant literature provides plausible explanations for the declining trend of inventor productivity. On the one hand, inventors have an incentive to stick to established technologies since they have accumulated expertise on them through learning-by-doing (Horii, 2012). On the other hand, once made public, technologies face incessant innovations by which new technologies replace old ones (e.g., Grossman and Helpman, 1991b; Klette and Kortum, 2004). Publicized technologies also attract imitations that deprive opportunities to profit by refining them (e.g., Chu, 2009; Cozzi and Galli, 2014). The latter negative effects eventually dominate, since

[1] For example, Romer (1990); Grossman and Helpman (1991a); Aghion and Howitt (1992); Kortum (1997); Klette and Kortum (2004); Acemoglu et al. (2017); Akcigit and Kerr (2018).

[2] For example, Grossman and Shapiro (1978); Chang (1995); Matutes et al. (1996); Schotchmer (1996); König et al. (2014); Panebianco et al. (2016).

[3] For example, Duranton and Puga (2001); Bettencourt et al. (2007); Davis and Dingel (2018).

[4] Somewhat informal studies can be found in, for example, Breschi et al. (2003); Garcia-Vega (2006); Østergaard et al. (2011); Huo and Motohashi (2015); Inoue et al. (2015); Akcigit et al. (2018).

[5] Since an applied patent does not necessarily result in an innovation, our analyses are essentially about knowledge creation rather than innovation, although we use these two terms interchangeably.

[6] Here, three five-year periods between 1995 and 2009 are considered.

learning-by-doing is subject to decreasing returns (Horii, 2012).

How do successful inventors stay productive in these circumstances? Horii (2012) proposed a model of innovation associated with technological shifts. In his model, consumers wish to satisfy an indefinite range of wants, which induces an inventor to seek an unexplored technological niche where he or she can create demand for new products realized by new technology. While his model lacks a micro mechanism behind the technological shifts, it is complemented by Berliant and Fujita (2008).[7]

In the Berliant-Fujita model, agents communicate via common knowledge and invent in pairs by utilizing their mutual differentiated knowledge, where an appropriate balance between common and differentiated knowledge facilitates collaborative innovation. A longer duration of collaboration by the same pair increases their common knowledge while decreasing their mutual differentiated knowledge, which at the same time accumulates differentiated knowledge between them and the remaining agents. To maintain the best knowledge composition, agents optimally choose the set of their collaborators and the allocation of time for each collaboration.

Given these facts and theoretical backgrounds, we develop three separate regression models. The first model represents the pairwise "knowledge creation function" proposed by Berliant and Fujita (2008). In this model, we focus on the differentiated knowledge of collaborators, as this is an obvious source of new ideas that would take an inventor to an unexplored technological niche. It is quantified in terms of the quality/novelty-adjusted output of the collaborator *excluding* the patents developed jointly with the inventor. We find that a 10% increases in collaborators' differentiated knowledge for an inventor raises his or her quality- and novelty-adjusted research output by around 2.8% and 3.5%, respectively, which thus implies positive but decreasing returns of this knowledge, consistent with the theoretical model.

In the second model, we decompose the contribution by collaborators' differentiated knowledge to the research output of an inventor (computed from the regression of the first model) into the fraction accruing to the quality/novelty and that to the quantity of his or her output. We find that the contribution is mostly dedicated to increasing the quantity, rather than the quality, of research output under the quality-adjusted productivity measure. But, as large as around 65% of the contribution accounts for increasing the novelty, rather than the quantity, of research output under the novelty-adjusted productivity measure. It follows that a major role of collaboration is to induce the technological shift of an inventor to a new niche, which is consistent with Berliant and Fujita (2008) as well as Horii (2012).

In the third model, we probe into the factors determining the amount of differen-

[7]The Berliant-Fujita model is to our knowledge the only explicit formalization of collaborative innovation at the individual inventor level. Weitzman (1998) and Olsson (2000, 2005) proposed formulations in which new ideas are generated by the recombinations of extant ideas. However, in their models, the process through which such recombinations take place are passive.

tiated knowledge that each inventor obtains from his or her collaborators. Here, we focus on the role of collaborator recombinations and find that a more active recombination has a selection effect in collaborations, resulting in a set of new collaborators with a larger average differentiated knowledge. We find that a 10% increase in the new collaborators of an inventor raises the average quality- and novelty-adjusted differentiated knowledge of collaborators by around 14% and 18%, respectively.

While we also find that more able collaborators tend to be attracted to the top inventors with a particularly large stock of knowledge, inventors with a smaller stock of knowledge may still be able to compensate for their shortage of knowledge by more active recombinations of collaborators. These findings explain the observed upgrading of inferior inventors as well as the persistent productivity of top inventors.

In these regressions, we control for individual fixed effects by exploiting panel data, and a variety of firm, industrial as well as other local factors. Yet, we face identification problems due to network endogeneity stemming from endogenous collaborations of inventors to maximize their productivities. The identification and estimation of models with endogenous networks are substantial challenges in the literature on network econometrics (e.g., Jackson et al., 2017). We argue, however, that each endogenous variable for an inventor in our models can be reasonably instrumented by the average value of the same variable for his or her distant indirect collaborators. Typically, adding degrees of separation in the network is double-edged, since it not only reduces the reflection problem but also makes the instrument weaker. But, we benefit from a special situation in which the relevance of the instrument is extrinsic to the inventor network as it comes from the assortative matching by productivity among firms and workers. The matching is essentially exogenous to individual inventors given that it takes place prior to the formation of research network, and is based on more diverse aspects than on R&D activities.[8] As a consequence, the relevance of the instruments is maintained even when the information of only distant indirect collaborators is used, as long as the assortative matching affects the indirect collaborators and the targeted inventors simultaneously.

The rest of the paper is organized as follows. We start by making key observations about the dynamics of knowledge creation and inventor productivities in Section 2. The related literature is reviewed in Section 3. The Berliant-Fujita model is described in Section 4 and the corresponding regression models are developed in Section 5. Data are detailed in Section 6, the identification strategy is discussed in Section 7, and the baseline regression results are presented in Section 8. A series of robustness checks are done in Section 9. Concluding remarks are made in Section 10.

[8] See, for example, Mori and Turrini (2005); Mendes et al. (2010); Bartolucci and Devicienti (2013); Behrens et al. (2014); Eeckhout and Kircher (2018); Gaubert (2018). In Section 7.3, we add supportive evidence from the financial and ownership data of firms in Japan.

2 Facts

To guide our analyses to follow, we make three observations on patent development in Japan, while postponing the description of the data to Section 6.

2.1 Productivity of an inventor

Our panel data consist of three periods, each of which aggregates five consecutive years: period 0 includes years from 1995 to 1999, period 1 from 2000 to 2004, and period 2 from 2005 to 2009. We focus on the balanced set I of 107,724 inventors, each of whom participated in at least one patent in each period.

Let $\mathcal{G}_{it}$ be the set of patents in which inventor i participates in period t, and G_j for $j \in \mathcal{G}_{it}$ be the set of inventors who participate in patent j. Denoting the output of patent project j by a scalar $g_j > 0$, *the productivity of inventor i* is defined in terms of the total output of patents in which he or she participated, with the output of each patent being discounted by the number of inventors involved in the patent:

$$\bar{y}_{it} = \sum_{j \in \mathcal{G}_{it}} g_j / |G_j| \tag{2.1}$$

where $|G_j|$ means the cardinality of set G_j. (Hereafter, the expression $|X|$ for any set X means the cardinality of X.)

We consider two aspects of inventor productivity. One is *quality* based on cited count, where g_j represents the count of citations that patent j received in three years of publication. The other is *novelty*, where g_j represents the *degree of technological novelty* of patent j defined by the reciprocal, $1/r_j$, of the order, $r_j = 1, 2, \ldots$, of j in terms of its application date among all the patents classified in the same technological category as j.[9,10] The technological category of a patent is identified by the "subgroup" of the International Patent Classification (IPC) in the baseline analyses.[11]

---

[9]Our data include all the patents applied in 1993 and thereafter as well as some older applications published in 1993 or later. Thus, by construction, our measure of novelty tends to overstate the novelty in technological categories defined before 1993. However, since our regression analyses use novelty data from 2000 and later (i.e., periods 1 and 2), the effect of truncation should not be too problematic as we have a seven-year lead time before 2000. The remaining overstatement is also controlled by the period fixed effect.

[10]Our novelty measure reflects *nicheness* of technological invention publicized by the patent. It can also be interpreted as an inverse measure of *crowdedness* in the market for the technological category.

[11]About 40,000 IPC subgroups are active in each period, and a single primary IPC subgroup is assigned to each patent. Refer to Section 6.1.2 for the details.

2.2 Dynamics of the relative productivities of inventors

This section discusses the dynamics of relative productivities of inventors. Let $I_t^{\mathrm{TOP}}(x)$ represent the set of inventors in the top x% in I in terms of their productivity in each period $t = 0, 1$, and 2. The set of inventors in each 5% interval of the productivity percentiles from 0% to 100% can then be expressed by $\Gamma_t(x) \equiv I_t^{\mathrm{TOP}}(x) \backslash I_t^{\mathrm{TOP}}(x-5)$ for $x = 5, 10, \ldots, 100$, where "\" is a set difference operator. Call $\Gamma_t(x)$ the (*productivity*) *class* x of inventors in period t.

For classes, $x = 5, 10, \ldots, 100$, under quality- and novelty-adjusted productivities, the height of each blue bar in Panels (a) and (b) in Figure 2.1, respectively indicates the share of inventors of class x in period 0 who stay at least in the same class x' ($\leq x$) in period 1. The graphs reveal a clear pattern:[12]

Observation 1 (Churning of relative productivities) *Under either measure of productivity, fewer than half of inventors above the median productivity $x < 50$ in period $t-1$ remain at least as productive in period $t \in \{1, 2\}$, indicating a strong pressure to prevent inventors from maintaining their relative productivity. In other words, a sizable proportion of inferior inventors replace superior ones in their productivity ranking in each period.*

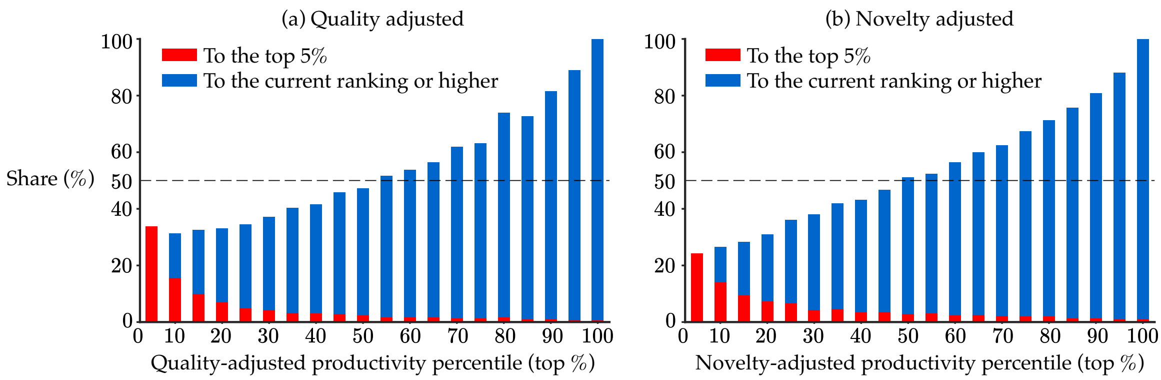


Figure 2.1: Change in the productivity class of inventors from period 0 to period 1

As discussed in the Introduction, a major reason for this downward pressure may be the obsolescence and imitations of technologies as well as decreasing returns in learning-by-doing from the extant technologies. Yet, we find that some top inventors stay highly productive, while some inferior ones surpass superior inventors. Each red bar in Figure 2.1 indicates the share of inventors in the corresponding class in period 0 who transitioned to the top 5% class in period 1. Although upgrading to the top 5% is less likely for inventors in a lower class, the transitions are observed from a wide range of lower classes.[13]

[12] A similar result is obtained for the transition from periods 1 to 2.

[13] A similar observation was made for US data between 1880 and 1940 by Akcigit et al. (2017), who found evidence that new inventors receive more patent citations than incumbent inventors.

2.3 Collaborator recombinations and technological shifts

We next present the key relationship among productivity, collaboration, and technological specialization of inventors that suggests the mechanism behind knowledge creation. Let

$$N_{it} \equiv \cup_{j \in \mathcal{G}_{it}} G_j \backslash \{i\} \tag{2.2}$$

represent the set of collaborators of inventor $i \in I$ in period t such that each inventor in N_{it} participates in the development of at least one common patent with i in period t. The *collaborator recombination* of inventor i in period t is then defined by

$$\Delta n_{it} \equiv |N_{it} \backslash N_{i,t-1}| \tag{2.3}$$

i.e., the number of new collaborators in period t.[14] The average values of Δn_{it} for inventors in I are 9.84 and 6.37 in periods 1 and 2, respectively. Provided that the number of collaborators is the same across periods, these values coincide with the average numbers of collaborators that were replaced.

Next, define the *technological specialization* of inventor i in period t by set S_{it} of the IPC subgroups associated with the patents in which inventor i is involved in period t. The *technological shift* of inventor i is then defined, similarly to the collaborator recombination in (2.3), by the number of IPC subgroups in which i is newly specialized in period t:

$$\Delta s_{it} \equiv |S_{it} \backslash S_{i,t-1}|. \tag{2.4}$$

The average values of Δs_{it} are 4.41 and 2.66 in periods 1 and 2, respectively. High correlations, 0.55 and 0.54, between $\ln \Delta n_{it}$ and $\ln \Delta s_{it}$ in periods 1 and 2, respectively suggest that new collaborations result in a shift in inventors' technological expertise.

For what purpose, do inventors shift their technological specialization? As discussed in the Introduction, Horii (2012) considered an economy in which demand for new technologies always exists, so that inventors have incentives to shift their technological expertise to unexplored niches and develop novel technologies. If the collaborator recombination is an effective means for this purpose as modeled by Berliant and Fujita (2008), other things being equal, he or she is more likely to achieve a technological niche (i.e., a larger novelty-adjusted $\bar{y}_{it}$) in the current period through the technological shift, Δs_{it}, realized by a larger collaborator recombination, Δn_{it}, from the previous period.

In the Berliant-Fujita model, technological shifts are realized by utilizing the

[14] Alternatively, it may be defined by the sum of the number of new collaborations and that of separations from the collaborations in the previous period, i.e., $\Delta n_{it} = |N_{it} \backslash N_{i,t-1}| + |N_{i,t-1} \backslash N_{it}|$. The qualitative result remains the same under both definitions.

differentiated knowledge of new collaborators. Our data are highly suggestive of this causality, as the correlations between novelty-adjusted $\ln \bar{y}_{it}$ and $\ln \Delta s_{it}$ are 0.30 and 0.29 in periods 1 and 2, respectively, in addition to the high correlation between $\ln \Delta s_{it}$ and $\ln \Delta n_{it}$ mentioned above.

These high correlations naturally extend to include the quality-adjusted productivity measure. To see this, consider the sets of inventors who stay in a given quality-adjusted productivity class $x = 5, 10, \ldots, 100$ persistently in both periods 1 and 2, i.e., $\Gamma(x) \equiv \cap_{t=1,2} \Gamma_t(x)$. Denote *the average collaborator recombination* by an inventor in class x in period t by

$$\Delta n_t(x) \equiv \frac{1}{|\Gamma(x)|} \sum_{i \in \Gamma(x)} \Delta n_{it} \tag{2.5}$$

the *average technological shift* by an inventor in class x in period t by

$$\Delta s_t(x) \equiv \frac{1}{|\Gamma(x)|} \sum_{i \in \Gamma(x)} \Delta s_{it} \tag{2.6}$$

and the *average productivity* of an inventor in class x in period t by

$$\bar{y}_t \equiv \frac{1}{|\Gamma(x)|} \sum_{i \in \Gamma(x)} \bar{y}_{it}. \tag{2.7}$$

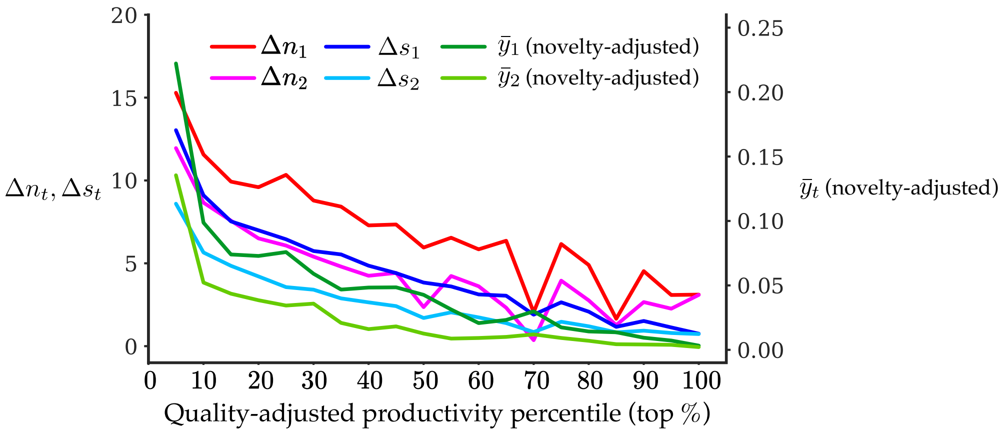


Figure 2.2: Recombinations, technological shifts, and productivities of inventors

Figure 2.2 plots $\Delta n_t(x)$, $\Delta s_t(x)$, and novelty-adjusted $\bar{y}_t$ for $t = 1, 2$ for each quality-adjusted productivity class $x = 5, 10, \ldots, 100$. There is a clear increasing tendency of all three measures for more quality-wise productive inventors (i.e., for a smaller x). Taken together, our observation can be summarized as follows.

Observation 2 (Recombinations, technological shifts, and inventor productivities) *A more quality-wise productive inventor practices a more active recombination of collaborators and is associated with a larger technological shift as well as higher novelty in the created knowledge on average.*

2.4 Invention strategies by productivity level

Our final observation is on the difference in the actions taken by inventors with different productivity levels. Panels (a) and (b) in Figure 2.3 show the distributions of collaborator recombinations and the novelty-adjusted productivity of inventors in period 1 for the top 10% and bottom 10% inventors under quality-adjusted productivity.[15,16] Both distributions are substantially right skewed for the top 10% inventors. That is, although both Δn_{it} and the novelty-adjusted $\bar{y}_{it}$ are larger on average for the top 10% than the bottom 10% inventors, a substantial population of the top 10% do not seek new collaborations or novelty in developed technologies.

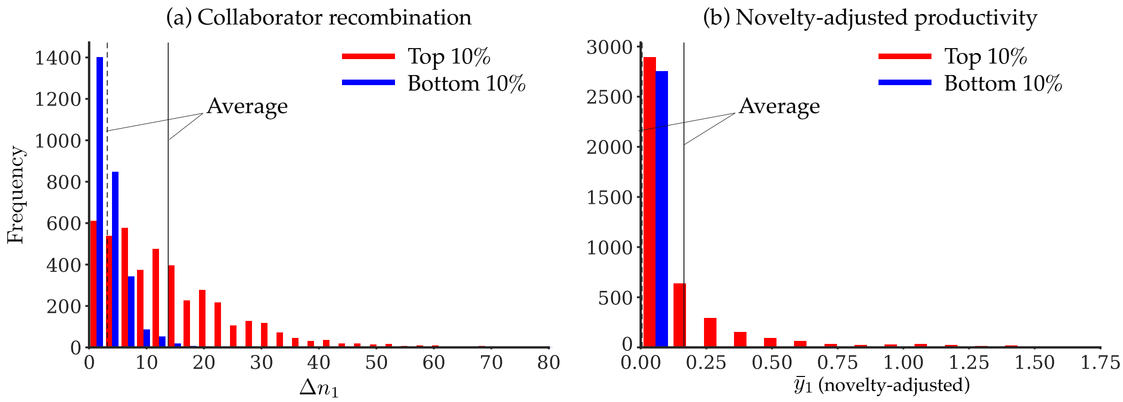


Figure 2.3: Collaborator recombinations and novelty-adjusted productivity of the top 10% and bottom 10% inventors quality-wise in period 1

On the one hand, the top 10% inventors appear to rely on their high innate ability and/or a large stock of knowledge to maintain their high quality-adjusted productivity without seeking novel technologies. On the other hand, in the context of the Berliant-Fujita model, the right skewness of the red plots in Figure 2.3 implies that an inventor without high innate ability or a large stock of knowledge may still be able to raise his or her productivity by finding new collaborators who have more relevant differentiated knowledge to enhance his or her technological expertise.

The difference in invention strategy between more and less established inventors is actually suggested by the data directly. To see this, let the size of knowledge stock of inventor i in period t be quantified by the cumulative number of technological categories, $k_{it} = \left|\bigcup_{t'<t} S_{it'}\right|$, that inventor i has worked on in the past.[17] Let $\Delta n_t^{\text{Top 5\%}}(x)$ represent the average size of collaborator recombinations by inventors who upgraded their productivity class from x in period $t-1$ to the top 5% in period t. Similarly, let $\Delta n_t^{\text{Down}}(x)$ be the average size of the collaborator recombinations of

[15]Namely $\Gamma(5)\cup\Gamma(10)$ and $\Gamma(95)\cup\Gamma(100)$, respectively

[16]Similar distributions are obtained for period 2.

[17]The top 5% inventors quality-wise have on average 3.3 and 2.3 times more stock of knowledge than the bottom 5% quality-wise in periods 1 and 2, respectively, so that established inventors can potentially rely more on their stock of knowledge to create new knowledge than less established ones.

inventors who downgraded their productivity class from x in period $t-1$ to $x' > x$ in period t for $x = 5, 10, \ldots, 95$.[18]

Panels (a) and (b) in Figure 2.4 plot the ratio of relative size of collaborator recombination, $\Delta n_t^{\text{Top 5\%}}(x)/\Delta n_t^{\text{Down}}(x)$, of up- to downgrading inventors for each productivity class x under quality- and novelty-adjusted measures of productivity, respectively.

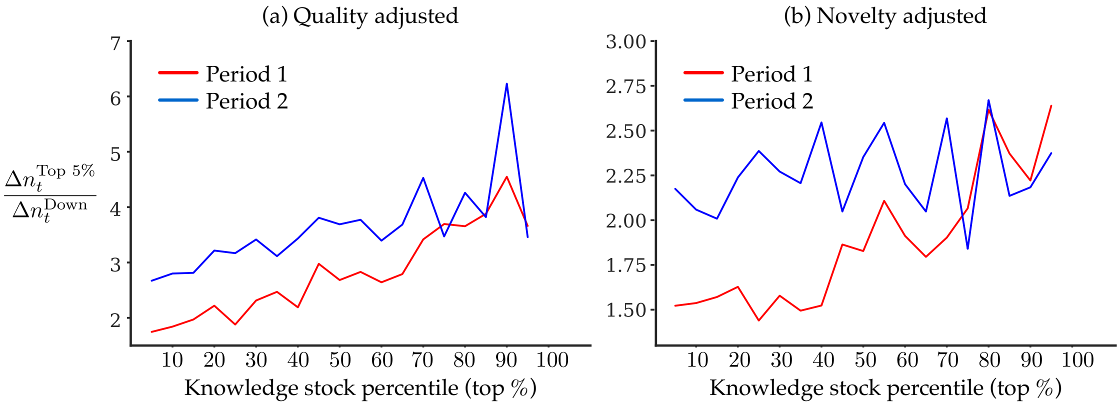


Figure 2.4: Recombination of up- versus downgrading inventors in period 1

Although the case of novelty-adjusted productivities in period 2 shown in Figure 2.4(b) is an exception, one can find a general tendency summarized as follows.

Observation 3 (Collaborator recombination versus stock of knowledge) *Inventors with a larger stock of knowledge rely relatively more on their own stock of knowledge than knowledge from new collaborators for inventing, while the opposite is true for inventors with lower stock of knowledge.*

3 Literature

The literature related to knowledge creation is diverse, including economic growth, industrial organization, and regional economics. This section provides an overview.

3.1 Theories

The first formalization of intentional innovation was in the public research sector by Shell (1966, 1967). A large variety of market-driven innovations by private sectors were proposed in the 1990s and thereafter (e.g., Romer, 1990; Grossman and Helpman, 1991b,a; Aghion and Howitt, 1992; Kortum, 1997), where investment decisions on R&D were explicitly modeled. The positive externalities from the accumulated human capital improved productivity in knowledge creation and drove economic growth in these models. Klette and Kortum (2004) linked innovation

[18]The lowest class $x = 100$ is omitted since there is no further downgrade from there.

technologies at the firm level to firm dynamics and then to growth at the economy level by extending the quality ladder-model by Grossman and Helpman (1991b), and a number of extensions of their model followed (e.g., Lentz and Mortensen, 2008; Akcigit and Kerr, 2018; Akcigit et al., 2016; Acemoglu et al., 2017). In their model, the product scope of a firm is interpreted as the stock of knowledge, which agrees with the argument by Weitzman (1998) that a new idea can generate a larger number of other new ideas if recombined with a larger number of existing ideas. These models, however, abstract from the mechanism through which the stock of knowledge is utilized by inventors in a firm to make innovations happen.[19]

The strategic aspects of innovation by individual firms have been explored by utilizing the techniques of industrial organization and network science. For example, König et al. (2014) formulated a trade-off between R&D collaborations and product market competition among firms when R&D investment of a firm reduces not only the production cost of this firm, but also those of collaborating competitors. Panebianco et al. (2016) formalized the mechanism of technology diffusion among firms by modeling the market for innovation and the timing of diffusion in a given network of firms. However, they still abstracted from the mechanism of knowledge creation as well as from the endogeneity of firm network in which R&D collaborations or innovation diffusion take place.[20,21]

The literature on knowledge creation at the inventor level is scarce. To our knowledge, Olsson (2000, 2005) was the first successful attempt to formalize the notion of cogitation by an inventor in developing new ideas. Knowledge creation in his models, however, is treated as passive, following a given stochastic process.

The work by Berliant and Fujita (2008) is, to our knowledge, the first to formalize active knowledge creation by individual inventors, with a focus on collaborative knowledge creation.[22,23] In their model, the steady-state size of a collaborating set of agents depends on the relative importance of common versus differentiated knowledge, where a larger size will result if differentiated knowledge is more appreciated.The typical steady state of their model, however, does not replicate the observed churning of relative productivity among inventors discussed in Section 2.

[19]As an exception, a recent contribution by Akcigit et al. (2018) extended quality ladder models of growth by introducing endogenous formation of a research team in which each ex ante homogeneous individual inventor faces an endogenous choice to become a team leader or a team member.

[20]See also Yang and Maskus (2001); Glass and Saggi (2002); Tanaka (2006); Tanaka et al. (2007) for related analyses in the context of economic growth.

[21]This topic has often been studied in conjunction with cost and benefit of the properties of a given patent system (e.g., Grossman and Shapiro, 1978; Chang, 1995; Matutes et al., 1996; Schotchmer, 1996).

[22]Jovanovic and Rob (1989) proposed a related search model in which collaborations with more able partners are more likely to result in the development of better knowledge.

[23]A variety of extensions of Berliant and Fujita (2008) have been proposed. Berliant and Fujita (2011) augmented their model with foresights and the possibility of learning from public knowledge; Berliant and Fujita (2012) introduced the variation in distance among agents; and Berliant and Mori (2017) allowed for heterogeneity in the innovation technology among inventors.

Horii (2012) complemented the Berliant-Fujita model in this aspect. In his model, consumers have an indefinite range of wants and thus demand for new technologies always exists. On the production side, learning-by-doing and local spillovers from the technological vicinity induce innovations to take place at discrete locations in the technological space. As the productivity improvement from learning-by-doing is subject to decreasing returns, there is an incentive to deviate from the extant technology and innovate in a distant unexplored niche in which a firm can rouse demand for the new technology and make a profit. The cost of the technological shift in this deviation is implicit in Horii (2012), and the collaborative knowledge creation in Berliant and Fujita (2008) complements this aspect in return by providing a micro mechanism for achieving the shift.

3.2 Empirics

A sizable literature on the effects of R&D investment on innovation, firm productivity, and economic growth in a country started in the 1960s (e.g., Griliches, 1964, 1979; Scherer, 1982; Coe and Helpman, 1995). In particular, after Kortum (1997) and Klette and Kortum (2004), studies began to structurally estimate the variations of their models by using firm-level micro data (e.g., Lentz and Mortensen, 2008; Akcigit et al., 2016; Acemoglu et al., 2017; Akcigit and Kerr, 2018). While these studies relate innovation behavior, size and growth of firms to aggregate economic growth, their models are not designed to disclose the innovation mechanism at the inventor level within a firm, and the innovation technologies are typically not estimated directly.

One exception is the work by Akcigit et al. (2018), who estimated a reduced-form model of team-level innovation similar to the knowledge creation function of the Berliant-Fujita model. Their key variable is the quantity and quality of the interactions within a team. The crucial difference from our approach is that their "interaction" effect comes from the past experience of the team leader, rather than from the his or her present collaborators (as in our case).

Another large strand of the literature is on knowledge spillover and diffusion (e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Murata et al., 2014; Kerr and Kominers, 2015). Its concern is on the distance and routes on which innovated technology and knowledge spread, not on how the knowledge is created.

Breschi et al. (2003); Garcia-Vega (2006); Østergaard et al. (2011) developed measures of common and differentiated technological knowledge relevant for innovation, showing that diversified knowledge as well as the mutual relatedness of knowledge within a firm and the innovation productivity of the firm are positively correlated. By using Japanese patent data similar to ours, Huo and Motohashi (2015) found a positive correlation between differentiated knowledge among inventors within a firm and their innovation productivities, whereas Inoue et al. (2015) focused on innovations by firm pairs, and argued that there is a decreasing return to common knowledge between collaborators. However, these studies ignore the endogeneity of collaborations, and thus the underlying causality is not clear.

Finally, this paper is also related to econometric identification and estimation in the context of a linear model in which some regressors are derived from endogenous network formation. The network endogeneity that arises in this study comes from inventors' strategic interactions to maximize their productivity by collaborations. Most common way to deal with network endogeneity is to consider network formation model to identify and estimate the parameters of interest (e.g., Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016; Comola and Prina, 2014; Li and Zhao, 2016; Patacchini et al., 2017; Johnsson and Moon, 2017).[24] However, this approach imposes parametric restrictions on the network formation model, and the estimation is biased when the model is misspecified. Since the model by Berliant and Fujita (2008) provides no simple econometric model of network formation as will be clear in Section 4, this traditional approach fails in our case. Thus, in contrast, in this paper we do not suppose any parametric model of network formation; instead, we propose an alternative approach to deal with endogenous regressors for an inventor by instrumental variables that are constructed from the information of his or her indirect collaborators.

4 The Berliant-Fujita model

This section provides a brief overview of the theoretical model of knowledge creation proposed by Berliant and Fujita (2008).

In a given period of time, each agent develops new knowledge either in isolation or by collaborating in pairs, building on the stock of knowledge accumulated in the past. Let I be the set of all the agents who engage in knowledge creation, where all agents are assumed to be symmetric. Let $\delta_{ij} \in [0,1]$ be the proportion of time that agent $i \in I$ allocates for the collaboration with $j \in I$. If agent i works in isolation (i.e., collaborates with him- or herself), his or her knowledge creation is subject to constant returns technology, given by

$$y_{ii} = \begin{cases} ak_{ii}, & \delta_{ii} \in [0,1] \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

where $a > 0$, k_{ii} is the knowledge stock of agent i, and y_{ii} is the output. If he or she

[24] Another typical approach assumes exogeneity of network structure (e.g., Bramoullé et al., 2009; Bramoullé and Fortin, 2010; Akcigit et al., 2018).

instead collaborates with agent $j (\neq i)$, the joint output, y_{ij}, of this collaboration is given by

$$y_{ij} = \begin{cases} b\left(k_{ij}^C\right)^{\theta}\left(k_{ij}^D\right)^{\frac{1-\theta}{2}}\left(k_{ji}^D\right)^{\frac{1-\theta}{2}}, & \delta_{ij} \in [0,1] \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

where $b > 0$, k_{ij}^C is the common knowledge between i and j, k_{ij}^D is the knowledge of agent i differentiated from that of j, and $\theta \in (0,1)$ is the relative importance of common knowledge.

All knowledges are symmetric, and the output from the collaboration of agents i and j becomes their common knowledge. Thus, the common knowledge between i and j increases as their collaboration lasts longer, while the differentiated knowledge between i with other agents also increases relative to their common knowledge. To achieve the best combination of common and differentiated knowledge with collaborators, agents collectively decide the group of collaborators, where each agent i optimally chooses δ_{ij} for each j of his or her collaborators.[25]

In this context, the steady-state group size that maximizes growth in the knowledge stock is given by $1+1/\theta$, and the time allocation for collaborations is given by $\delta_{ij} \equiv \delta = 1/(1+1/\theta)$ for all $i, j \in I$.[26]

5 Regression model

This section introduces three regression models to identify the causal relationship among the quality/novelty of inventions, collaborators' differentiated knowledge, and magnitude of collaborator recombination at the inventor level based on the Berliant-Fujita model. In the regressions, we focus on collaborative inventions, and do not address the choice between working in collaboration and working in isolation. In other words, our formulation assumes a strictly positive number of collaborators for each inventor in each period.

Let $t = 0, 1, \ldots, T$ be the consecutive periods in which data are available and let I_t be the set of all inventors who participated in the development of at least one patent in period t. The subset of inventors, each of whom is involved in the development of at least one patent in every period (introduced in Section 2.1), is denoted by $I (\subset I_t)$.

Let $\mathcal{G}_t$ represent the set of all patents applied in period t. We call the development of each patent $j \in \mathcal{G}_t$ a *project* j. Then, G_j introduced in Section 2.1 represents the set of inventors who participated in project j, and the set of projects in which inventor $i \in I_t$ participated (also introduced in Section 2.1) can be rewritten as $\mathcal{G}_{it} \equiv \{j \in \mathcal{G}_t : i \in G_j\}$.

[25] Myopic core is adopted as the equilibrium concept.

[26] This is the steady-state equilibrium when agents initially have sufficient common knowledge, which is a natural situation for collaborations to start (Berliant and Fujita, 2008, Proposition 1).

Accordingly, set N_{it} of the collaborators of inventor i in period t is given by (2.2) in Section 2.3, and output, $\bar{y}_{it}$, of inventor i is given by (2.1) in Section 2.1.

5.1 Knowledge creation function

To bring the knowledge creation function (4.2) to the data, we modify the original specification. First, while it is defined for each of multiple pairwise collaborations,we formulate a regression model for a single *average pairwise knowledge creation function*:

$$\ln y_{it} = \alpha + \beta \ln k_{it}^{D} + \gamma_1 \ln k_{it} + \gamma_2 (\ln k_{it})^2 + \ln A_{it} + \lambda_i + \tau_t + \varepsilon_{it} \tag{5.1}$$

in which y_{it} represents the *average pairwise output* by inventor i:

$$y_{it} = \bar{y}_{it}/n_{it} \tag{5.2}$$

where $n_{it} \equiv |N_{it}|$.[27] The variation in pairwise productivities for a given inventor is assumed to be random and captured by inventor- and period-specific error term, ε_{it}.

Second, in (5.1), we focus on the differentiated knowledge, k_{ji}^{D}, of collaborators in (4.2), since this is a source of new ideas as discussed in Section 2, while abstracting from the role of common knowledge, k_{ij}^{C}, and that of differentiated knowledge, k_{ij}^{D}, of inventor i him- or herself. This key variable appears as k_{it}^{D} in the second term on the right-hand side (RHS) of (5.1) in the form of *the average pairwise differentiated knowledge of collaborators of* i, and is defined by the average output that the collaborators of i produced outside the joint projects with i:

$$k_{it}^{D} = \frac{1}{n_{it}} \sum_{j \in N_{it}} \sum_{k \in \mathcal{G}_{jt} \setminus \mathcal{G}_{it}} \frac{g_k}{|G_k|}. \tag{5.3}$$

Here, k_{it}^{D} includes only the *fresh* knowledge of collaborators that they create with inventors other than i in the current period and not their knowledge stock from the past. This definition reflects Observation 1 in Section 2.2 that past knowledge is strongly associated with negative effects. The value of k_{it}^{D} may also be interpreted as the average productivity of i's collaborators outside the joint projects with i. This feature plays a role when we construct an instrument for this variable in Section 7.

Third, as for the common knowledge k_{ij}^{C} and differentiated knowledge k_{ij}^{D} of inventor i in (4.2), their effects are controlled for by his or her stock of knowledge:

$$k_{it} = \Big| \bigcup_{t' < t} S_{it'} \Big|. \tag{5.4}$$

While this approach does not capture the roles of the common knowledge between

[27] Refer to Appendix A for the interpretation of average pairwise productivity.

i and his or her collaborators and of the differentiated knowledge of i precisely,[28] they are by definition expected to be positively correlated with the knowledge stock of i.[29] Moreover, the size of the knowledge stock is expected to control for a variety of other effects, including learning-by-doing as well as imitations and obsolescence effects on the extant technologies discussed in Section 2.2. We capture their overall effects up to the second order, the fourth term on the RHS of (5.1).

Finally, in the fifth term, A_{it} bundles the inventor- and time-specific productivity shifters for inventor i:

$$A_{it} \equiv e^{\boldsymbol{X}'_{it}\boldsymbol{\eta}}, \tag{5.5}$$

where $\boldsymbol{X}_{it}$ represents a vector including spillover effects from other inventors in the geographical neighborhood, proximity to R&D expenditure, manufacturing employment/production, and residential population.

The last three terms, λ_i, τ_t, and ε_{it}, on the RHS are the time-invariant inventor fixed effect, period fixed effect, and inventor- and period-specific error, respectively. The values of parameters $\alpha, \beta, \gamma_1, \gamma_2, \boldsymbol{\eta}$, and τ_t are estimated by regressions.

5.2 Quality/novelty and quantity decomposition

The definition of quality and novelty of output by an inventor given by (2.1) implies the log-linear relationship between quality and quality/novelty of his or her output:

$$\ln y_{it} = \ln y^p_{it} + \ln y^q_{it}. \tag{5.6}$$

In the first term on the RHS of (5.6), y^p_{it} denotes the quantity, i.e., the average count of patents, of inventor i's pairwise output given by

$$y^p_{it} \equiv \bar{y}^p_{it}/n_{it} \tag{5.7}$$

where $\bar{y}^p_{it} \equiv \sum_{j\in\mathcal{G}_{it}} 1/|G_j|$ which coincides with $\bar{y}_{it}$ under $g_j = 1$ in (2.1) ; whereas in the second term, y^q_{it} represents the average quality/novelty of i's pairwise output:

$$y^q_{it} \equiv y_{it}/y^p_{it} \left(= \bar{y}_{it}/\bar{y}^p_{it} \right). \tag{5.8}$$

[28] Ideally, the knowledge stock may be defined in terms of productivities as in the case of k^D_{it} in (5.3). However, this leads to an identification problem because of the endogeneity induced by including the lagged outcomes. Moreover, as shown in Section 2.2, publicized technologies are quickly imitated and thus become obsolete. Hence, the output measures may be unsuitable for defining the knowledge stock.

[29] In principle, it is possible to define average pairwise common knowledge and average pairwise differentiated knowledge of inventor i in terms of the technological categories as in knowledge stock, k_{it}. But, since these should strongly correlate with k_{it} by construction, and hence can be controlled for by k_{it} to a large extent, we focus on the identification of the most interested causal relationship between k^D_{it} and y_{it}, rather than allowing for multiple endogenous variables.

We can thus decompose the effect of each explanatory variable in (5.1) into those on the quantity and on the quality/novelty of inventors' pairwise output, y_{it}, by estimating the model given by

$$\ln y_{it}^{m} = \alpha^{m} + \beta^{m} \ln k_{it}^{D} + \gamma_{1}^{m} \ln k_{it} + \gamma_{2}^{m} (\ln k_{it})^{2} + \ln A_{it}^{m} + \lambda_{i}^{m} + \tau_{t}^{m} + \varepsilon_{it}^{m} \tag{5.9}$$

for $m = p$ and q, where the coefficients of each explanatory variable for $m = p$ and q add up to that of the corresponding variable in (5.1). In particular, we have $\beta = \beta^{p} + \beta^{q}$ for the effect of collaborators' differentiated knowledge.

5.3 Recombinations and differentiated knowledge of collaborators

Finally, we introduce our third regression model (5.10) to identify the factors determining the value of k_{it}^{D} in (5.1):

$$\ln k_{it}^{D} = \tilde{\alpha} + \tilde{\beta} \ln \Delta n_{it} + \tilde{\gamma}_{1} \ln k_{it} + \tilde{\gamma}_{2} (\ln k_{it})^{2} + \ln \tilde{A}_{it} + \tilde{\lambda}_{i} + \tilde{\tau}_{t} + \epsilon_{it} \tag{5.10}$$

where Δn_{it} given by (2.3) is considered to be endogenous as it is a result of the active efforts and/or random factors that influence the recombination of collaborators at the individual or firm/establishment level.

The aim of this regression is twofold. One is to see if the more substantial recombination of collaborators results in acquiring knowledge associated with higher quality/novelty from collaborators, as we interpreted Observation 2 in the context of the Berliant-Fujita model. The other is to see if the substitutability between the stock of knowledge and collaborator recombinations suggested by Observation 3 is relevant in raising the quality/novelty of collaborators' differentiated knowledge.

It is to be noted that we present the results of two separate estimations for (5.1) and (5.10), rather than incorporating the collaborator recombination explicitly in the knowledge creation function given by (5.1). The practical reason for the separation is the estimation problem in combining the two models (refer to footnote 45 in Section 8). But, it is also true that only the former has a specific microeconomic foundation according to Berliant and Fujita (2008). In the Berliant-Fujita model, each inventor optimally chooses the size of collaborator recombination to balance the common and differentiated knowledge between him or her and his or her collaborators, not just to maximize the value of the differentiated knowledge of collaborators. Thus, (5.10) captures only a part of the entire causality behind the determination of k_{it}^{D}. To be consistent with the Berliant-Fujita model, however, we expect that the overall effect of Δn_{it} through k_{it}^{D} on y_{it} exhibit positive but decreasing returns, since successful collaborations requires a certain share of common knowledge.

6 Data

In this section, we describe our dataset, focusing primarily on patent data.

6.1 Patent data

The patent data are taken from the *published unexamined patent applications* of Japan (Artificial Life Laboratory, 2018) which provides information on the published patent applications to be examined for approval rather than approved patents. The advantage of using unexamined applications is that their flow at a given point in time reflects the amount of research activities at that point more precisely than the flow based on approved patents. In this data, each inventor is uniquely identified as long as his or her name and affiliation have not changed.

6.1.1 Patent projects

Our analysis targets inventors who participated in the development of patents applied between 1995 and 2009 in Japan. Since a patent development is a time-consuming project whose underlying research could take several years, the productivity of an inventor is evaluated by his or her output over five years. The choice of a five-year window also reflects the availability of other relevant data from census.

We construct a three-period panel in which period 0 consists of years from 1995 to 1999, period 1 from 2000 to 2004, and period 2 from 2005 to 2009. Since k_{it} and Δn_{it} require information from the previous period, period 0 is not included in the regressions. The information in 2010–2016 is used to account for the time lag between the date of application and that of publication as well as to count the forward citations for each patent. Consequently, our panel for regressions consists of two periods, 1 and 2.

Table 6.1 summarizes the basic data.[30] In particular, we focus on the ($|I| =$) 107,724 inventors who have been active in all the three periods, although the information on other inventors is still used as long as they collaborated with the selected inventors.

The average number of inventors in a project throughout the study period is about two (row 7), hence is consistent with the assumption of pairwise collaboration in the Berliant-Fujita model. Since collaborations are typically polyadic: an inventor has six to nine collaborators on average (row 9), which also agrees with the implication

[30]The fact that the number of patents per inventor is declining over time may reflect the influence of the tendencies of block patents (e.g., Chu, 2009; Cozzi and Galli, 2014; Nicholas, 2014; Jell et al., 2017). In particular, after the applied unexamined patents were made public in digitized form in 1993, firms have stronger incentives to block potential competitors from innovating in their common technologies.

from the Berliant-Fujita model. Moreover, about 90% of inventors have at least one collaborator (row 8), which justifies our focus on collaborative knowledge creation.

Table 6.1: Descriptive statistics of basic variables

		Period	
		(1)	(2)
Variable		1	2
(1) Number of patents	$\left\lvert \bigcup_{i \in I} \mathcal{G}_{it} \right\rvert$	1,758,780	1,546,596
(2) Number of IPC classes		120	122
(3) Number of IPC subclasses		608	615
(4) Number of IPC subgroups	$\left\lvert \bigcup_{i \in I} S_{it} \right\rvert$	40,691	38,339
(5) Number of inventors in period t	$\lvert I_t \rvert$	1,208,197	1,094,789
(6) Number of inventors active in all periods	$\lvert I \rvert$	107,724	107,724
(7) Number of inventors per patent	$\lvert G_{jt} \rvert$	2.193	2.244
		(1.538)	(1.609)
(8) Share of collaborating inventors	$\left\lvert \{i \in I_t : \lvert N_{it} \rvert > 0\} \right\rvert / \lvert I_t \rvert$	0.896	0.868
(9) Number of collaborators per inventor	$\lvert N_{it} \rvert$	8.518	6.323
		(9.321)	(7.579)
(10) Number of new collaborators per inventor	Δn_{it}	6.893	4.354
		(7.907)	(5.848)
(11) Number of patents per inventor	$\lvert \mathcal{G}_{it} \rvert$	10.66	6.858
		(16.21)	(11.95)
(12) Number of IPC sections per inventor		1.812	1.533
		(0.952)	(0.799)
(13) Number of IPC classes per inventor		2.473	1.918
		(1.788)	(1.381)
(14) Number of IPC subclasses per inventor		2.984	2.241
		(2.409)	(1.874)
(15) Number of IPC subgroups per inventor	$\lvert S_{it} \rvert$	5.471	3.713
		(5.223)	(4.026)
(16) Size of cumulative IPC subgroups per inventor	$\lvert \cup_{t' < t} S_{it'} \rvert$	4.550	8.958
		(4.659)	(7.582)

Numbers in parentheses are standard deviations.

6.1.2 IPC

Each published patent application is associated with at least one technological classification based on the IPC, which is maintained by the World Intellectual Property Organization.[31] The IPC classifies technologies into eight sections: A (human necessities), B (performing operations; transporting),…, H (electricity). These sections are divided into classes such as A01 (agriculture; forestry; animal husbandry; hunting; trapping; fishing) and then into subclasses such as A01C (planting; sowing; fertilizing). Each subclass is further divided into groups, e.g., A01C1 (apparatus, or methods of use thereof, for testing or treating seed, roots, or the like, prior to sowing or planting), and then into subgroups, e.g., A01C 1/06 (coating or dressing seed) and A01C 1/08 (immunizing seed). The IPC's labeling scheme is consistent over time, and a newly introduced category is basically associated with a new technology (e.g., the classes B81 for microtechnology and B82 for nanotechnology introduced in 2000). Taken together, the set of technological categories specified in the IPC at a given point

[31] Website: http://www.wipo.int/portal/en/index.html.

in time roughly represents the set of the state-of-the-art technologies at that time, and hence makes an appropriate proxy for the set of technological knowledge.

Although an applicant can claim more than one IPC categories for his or her patent, we adopt only the primary IPC category of each patent to represent its technology in order to avoid subjective variation. Consequently, we have 121, 609, and 40,691 (123, 616, and 38,339) relevant IPC classes, subclasses, and subgroups, respectively for period 1 (period 2), associated with the applied patents in our data.

Let S denote the set of all the technological categories (in terms of either one of IPC class, subclass, or subgroup) and the technological category assigned to patent j be $s_j \in S$. The *technological specialization* of inventor i is then defined by

$$S_{it} = \cup_{j \in \mathcal{G}_{it}} \{s_j\}. \tag{6.1}$$

In the baseline regressions, we adopt IPC subgroups to construct S_{it} and quantify the cumulative stock of knowledge, k_{it}, defined by (5.4).

Similarly, to quantify the technological novelty defined in Section 2.1, we adopt IPC subgroups as they exhibit the largest variation among inventors. We also control for the IPC class fixed effect to account for the possible incompatibility of the quality/novelty adjustment of patents across different technology categories, where each inventor is associated with his or her most frequently engaging IPC class.

6.2 Productivity and differentiated knowledge

Table 6.2 lists the descriptive statistics for productivity variables. Our preferred measure of the quality of a given patent is the count of forward citations following Trajitenberg (2002); Akcigit et al. (2018).[32] In our baseline analysis, we count the forward citations of each patent within three years of the publication date following Akcigit et al. (2018) (row 1, columns 1 and 2).[33,34] In our data, the cited counts in the first three years from publication account for more than 75% of the total cited count in the first 10 years for all samples. Thus, using the three-year window to evaluate the patent quality appears to be reasonable.

[32]Cited counts may not be an optimal measure of patent quality when there is an incentive to block follow-up patents as discussed by Abrams et al. (2013).

[33]We also conduct the same analysis under the count of forward citations within five years of publication to check the robustness (see Section 9.3.1).

[34]It is assumed that there is at least one (self-)citation, namely $g_j \geq 1$, under the quality-adjusted measure. That is, the cited count for each patent is inflated by 1 if there is no self-citation to avoid dropping patents without citations. Some authors (e.g., Inoue et al., 2015) argue that the citation-adjusted output of a patent project should exclude self-citations by inventors in the project. Our analyses, however, include them since there is no clear incentive to inflate the cited counts for patents (unlike the case of academic papers); hence, the self-citations tend to reflect genuine technological dependence. In fact, we find no qualitative difference between the results with and without citation weights (see Section 9.3.1).

Alternatively, we use technological novelty based on the IPC subgroups introduced in Section 2. Since cultivating a novel technology requires knowledge, there may be a more direct relationship between knowledge input and technological novelty (row 1, columns 3 and 4).

Table 6.2: Descriptive statistics of productivity variables

Productivity measure		**Cited counts**		**Novelty**	
		(1)	(2)	(3)	(4)
Period		1	2	1	2
(1) Output of a patent	g_{jt}	1.535 (2.527)	1.423 (3.850)	0.013 (0.056)	0.009 (0.049)
(2) Productivity of an inventor	$\bar{y}_{it}$	7.906 (16.83)	5.048 (163.31)	0.047 (0.134)	0.024 (0.084)
(3) Pairwise productivity of an inventor	y_{it}	1.389 (3.160)	1.728 (175.04)	0.009 (0.049)	0.006 (0.032)
(4) Avg. diff. knowledge of collaborators	k^D_{it}	1.411 (7.520)	1.053 (4.539)	0.008 (0.043)	0.005 (0.034)

Numbers in parentheses are standard deviations.

As defined in (2.1), when the output of a project is attributed to each inventor, it is discounted by the number of inventors involved in the project. Row 2 of Table 6.2 lists the output, $\bar{y}_{it}$, attributed to inventor i under each alternative productivity measure. Row 3 of Table 6.2 lists the average pairwise productivity, y_{it}, of each inventor defined by (5.2), which corresponds to the output of the knowledge creation function (4.2) proposed by Berliant and Fujita (2008) under these alternative productivity measures. Finally, row 4 of Table 6.2 lists the average differentiated knowledge of collaborators defined by (5.3) for each productivity measure.

6.3 Locational factors

Possible influence of various exogenous locational factors on productivity in knowledge creation has been suggested by existing studies. We briefly describe each factor included in the regression, with the precise definitions relegated to Appendix B.

R&D activities are disproportionately concentrated in large cities (see Figure B.1(b) in Appendix I). If an *urban agglomeration* (UA) is defined as a contiguous area of population density at least 1000/km^2 with the total population at least 10,000,[35] in 2000, 99% of all inventors concentrated in UAs, 81% in the largest three UAs, and 54% in the largest UA (Tokyo), while the corresponding shares for population are 75%, 54% and 32%, respectively. Inventors located within a 10km buffer of any of all the 453 UAs are assigned to the closest UA, while otherwise their locations are considered to be rural. In the regressions, the standard errors are clustered by UAs.[36]

[35]The population data are obtained from the Population Census (2010a) by MIAC.

[36]As UAs on average expand spatially over time, we used the boundaries of UAs in 2010, each of which provides on average the largest spatial extent during the study period of 1995–2009. However,

Local concentrations of four types of activities that potentially influence inventor productivity are considered: the concentrations of inventors (a_{it}^{INV}), R&D expenditure ($a_{it}^{\mathrm{R\&D}}$), manufacturing employment ($a_{it}^{\mathrm{MNF}_e}$) and output ($a_{it}^{\mathrm{MNF}_o}$). These partly account for the factors specific to the firm and establishment that an inventor belongs to. We also control for residential population (a_{it}^{POP}). Each local concentration is defined by the size of concentration in a circle of given radius around inventor i.

7 Identification by instrumental variables

This section presents our strategy for identifying the causalities behind knowledge creation by dealing with the endogeneity of the differentiated knowledge and recombinations of collaborators for individual inventors. There are two sources of endogeneity. One comes from inventors' endogenous collaboration, i.e., network endogeneity, where there exist unobservables which influence inventors' collaboration decisions as well as their productivities. The other comes from the mutual dependence of productivities between an inventor and his or her collaborators through k_{it}^{D} in model (5.1) (as well as (5.9)). This is the so-called reflection problem in the context of econometric network analysis (e.g., Manski, 1993; Bramoullé et al., 2009). In our case, however, we argue that the endogenous variables, k_{it}^{D} in model (5.1) and Δn_{it} in model (5.10), for inventor i can be instrumented by the average value of the same variable for the indirect collaborators of i.

Below, we formally define the instruments for the endogenous variables in Section 7.1, and explain their exogeneity and relevance in Sections 7.2 and 7.3, respectively. (In Appendix C, we briefly discuss the similarity and difference in the nature of endogeneity and the approach to get around the issue between the linear-in-means models of social interactions as in Bramoullé et al. (2009) and our model.)

7.1 Instruments

Let $\bar{N}_{it}^{\ell}$ be the set of up to the ℓ-th indirect collaborators of inventor i given by

$$\bar{N}_{it}^{\ell} = \bar{N}_{it}^{\ell-1} \cup \left[\bigcup\nolimits_{j\in \bar{N}_{it}^{\ell-1}} N_{jt} \right] \quad \ell = 1,2,\ldots \tag{7.1}$$

where the set of the "0-th indirect collaborators" is defined by the set of inventors consisting of i and his or her direct collaborators, $\bar{N}_{it}^{0} \equiv N_{it} \cup \{i\}$. To obtain $\bar{N}_{it}^{\ell}$ from $\bar{N}_{it}^{\ell-1}$ for each $\ell = 1,2,\ldots$, we expand $\bar{N}_{it}^{\ell-1}$ by the union of all the direct collaborators

the choice of the particular time point should not affect the basic results since most inventors are concentrated in relatively large UAs whose spatial coverage is relatively stable over the study period.

of $j \in \bar{N}_{it}^{\ell-1}$ as in (7.1). The set of the *ℓ-th indirect collaborators* of i can then be given by

$$N_{it}^{\ell} = \bar{N}_{it}^{\ell} \backslash \bar{N}_{it}^{\ell-1} \quad l = 1, 2, \ldots \tag{7.2}$$

The instruments, $k_{it}^{\mathrm{IV}_\ell}$ for k_{it}^{D} and $\Delta n_{it}^{\mathrm{IV}_\ell}$ for Δn_{it}, are constructed as the average value of the differentiated knowledge of collaborators and that of collaborator recombination, respectively for each ℓ-th indirect collaborator $j \in N_{it}^{\ell}$:

$$k_{it}^{D,\mathrm{IV}_\ell} = \frac{1}{n_{it}^{\ell}} \sum_{j \in N_{it}^{\ell}} k_{jt}^{D} \quad \text{and} \quad \Delta n_{it}^{\mathrm{IV}_\ell} = \frac{1}{n_{it}^{\ell}} \sum_{j \in N_{it}^{\ell}} \Delta n_{jt}. \tag{7.3}$$

Indirect collaborators may be weighted by the frequency of their appearance:

$$k_{it}^{D,\mathrm{IV}_\ell} = \frac{1}{\tilde{n}_{it}^{\ell}} \sum_{l \in \bar{N}_{it}^{\ell-1}} \sum_{j \in N_l} k_{jt}^{D} \quad \text{and} \quad \Delta n_{it}^{\mathrm{IV}_\ell} = \frac{1}{\tilde{n}_{it}^{\ell}} \sum_{l \in \bar{N}_{it}^{\ell-1}} \sum_{j \in N_l} \Delta n_{jt} \tag{7.4}$$

where $\tilde{n}_{it}^{\ell} \equiv \sum_{j \in \bar{N}_{it}^{\ell-1}} n_j$. Inventor j may appear more than once in the construction of the instruments in (7.4). Weighting by appearance frequency of indirect collaborators in the linkage helps strengthen the relevance of the instruments.

7.2 Exogeneity

This section explains how our instruments can virtually eliminate the endogeneities caused by the reflection problem and inventors' unobserved variables.

7.2.1 Reflection problem

The existing models of social interactions (e.g., Bramoullé et al., 2009; De Giorgi et al., 2010; Calvó-Armengol et al., 2009) suggest two reasons that the reflection effects in our context can be reduced by using instruments constructed from more distant indirect collaborators. One is the *distance effect* that the farther an indirect collaborator is from an inventor in the collaboration network, the smaller is the influence of the indirect collaborator's output on the inventor's output through the simultaneous equations of inventions and the indirect connections on the network.[37] Thus, by constructing instruments from sufficiently distant indirected collaborators, the reflection effects can be virtually eliminated. The other is the *averaging effect*. As long as the number of ℓ-th indirect collaborators increases as ℓ increases, the reflection effect on an inventor from each of his or her ℓ-th indirect collaborators is mitigated by averaging over a larger number of indirect collaborators.

[37] For example, in eq. (6) of Bramoullé et al. (2009), the endogenous peer effect from the ℓ-th indirect peer is given by $\beta^{1+\ell}$, where $\beta \in (0,1)$ and $\ell = 0, 1, 2, \ldots$ with the 0-th indirect peer being the direct peer. The indirect peer effects, $\beta^{1+\ell}$, from the ℓ-th indirect peers diminishes as ℓ increases.

Fortunately, the research network in our data consists of a set of large network components, so that we could identify up to fifth-indirect collaborators for a large number of inventors. Column 1 of Table 7.1 lists the average number of the ℓ-th indirect collaborators of an inventor for $\ell = 0$ to 5, where the 0-th indirect collaborators are the direct collaborators. The number of indirect collaborators of an inventor increases dramatically from 8.52 to 4,251 (6.32 to 2,563) for $\ell = 0$ to 5 in period 1 (period 2),[38] suggesting that the reflection emanating from each fifth indirect collaborator have only marginal effect due to both distance and averaging effects.

7.2.2 Unobserved factors

We suppose that inventors with similar (observable and unobservable) characteristics have proclivities to collaborate with each other, hence they might also have influence on their mutual productivities, while we also suppose that more distant indirect collaborators share less common characteristics with each other. Thus, by constructing instruments from sufficiently distant indirect collaborators, we are supposed to be able to eliminate the effects from unobserved factors.

The most plausible situation in which unobserved factors become problematic perhaps is the case that inventors have similar technological specialization. These inventors likely share opportunities and environment to exchange and learn ideas from each other, for example, through seminars, conferences and journals of common research subjects, which in turn affect their R&D productivities. Our data indicate, however, that the commonality of research subjects between a pair of inventors diminishes rapidly and eventually becomes negligible as the degrees of separation in the network increases between the pair.

Columns 2-8 in Table 7.1 list the diversity, $|S_{it}|$, of technological specialization of indirect collaborators in terms of IPC sections, classes, subclasses and subgroups, respectively. While an inventor on average specializes in 1.81, 2.47, 2.98 and 5.47 (1.53, 1.92, 2.24 and 3.71) in these categories, respectively in period 1 (period 2) (rows 12-15 in Table 6.1), the diversity in technological specialization increases for more indirect collaborators. For their fifth indirect collaborators, these numbers reach 7.50, 83.07, 275.8 and 1400 (7.29, 70.14, 213.3 and 1007), respectively in period 1 (period 2) (rows 6 and 12 of columns 2-5 in Table 7.1). Since the total number of IPC sections is eight, they are almost fully covered. For IPC classes, subclasses and subgroups, they cover 98.8%, 96.9% and 64.6% (97.9%, 94.2% and 60.6%), respectively among all the patents applied in period 1 (period 2). The set of the fifth-indirect collaborators of an inventor thus consists of almost all the kinds of specialists.

[38] It is also to be noted that the average number of the ℓ-th indirect collaborators more than doubles that of $(\ell-1)$-th indirect collaborators up to $\ell = 5$. Thus, the fifth-indirect collaborators of an inventor are still confined within a close neighborhood of the inventor relatively to the entire network.

Table 7.1: Diversity and similarity of technological specialization of inventors

Indirectness ℓ	(1) Count	(2) Diversity in technological category $\lvert S_{it}\rvert$ Section	(3) Class	(4) Subclass	(5) Subgroup	(6) Similarity in technological category j_{it}^{ℓ} Section	(7) Class	(8) Subclass	(9) Subgroup
	Period 1								
(1) 0	8.518	3.207	7.214	10.88	27.17	0.709	0.569	0.499	0.330
	(9.321)	(1.681)	(6.313)	(10.97)	(29.65)	(0.197)	(0.229)	(0.236)	(0.214)
(2) 1	51.36	4.670	17.38	32.49	94.83	0.567	0.370	0.280	0.087
	(64.79)	(2.057)	(14.62)	(32.57)	(102.82)	(0.215)	(0.228)	(0.220)	(0.120)
(3) 2	205.41	5.860	32.48	72.72	242.03	0.493	0.280	0.193	0.045
	(284.89)	(2.074)	(23.65	(64.70)	(240.91)	(0.202)	(0.202)	(0.186)	(0.083)
(4) 3	665.27	6.691	50.07	129.69	492.17	0.433	0.213	0.135	0.025
	(944.31)	(1.869)	(30.23)	(97.96)	(433.84)	(0.188)	(0.176)	(0.155)	(0.062)
(5) 4	1794.44	7.200	67.56	199.24	866.81	0.380	0.161	0.092	0.014
	(2385.52)	(1.584)	(33.33)	(126.37)	(675.22)	(0.172)	(0.148)	(0.123)	(0.046)
(6) 5	4250.87	7.501	83.07	275.77	1399.86	0.341	0.125	0.064	0.009
	(5076.27)	(1.314)	(33.31)	(146.28)	(969.11)	(0.156)	(0.124)	(0.098)	(0.036)
	Period 2								
(7) 0	6.323	2.658	5.057	7.352	17.57	0.757	0.648	0.588	0.432
	(7.579)	(1.530)	(4.642)	(7.972)	(21.84)	(0.207)	(0.247)	(0.263)	(0.271)
(8) 1	36.79	4.073	12.35	22.07	63.30	0.582	0.404	0.312	0.100
	(48.06)	(2.006)	(10.85)	(23.28)	(74.30)	(0.246)	(0.264)	(0.258)	(0.149)
(9) 2	137.59	5.306	23.84	50.43	164.62	0.505	0.309	0.218	0.054
	(195.61)	(2.147)	(18.70)	(48.17)	(179.36)	(0.229)	(0.234)	(0.219)	(0.106)
(10) 3	424.14	6.256	38.60	93.40	343.90	0.443	0.237	0.153	0.030
	(642.62)	(2.022)	(25.89)	(77.71)	(341.13)	(0.210)	(0.204)	(0.182)	(0.077)
(11) 4	1115.16	6.888	54.59	148.78	617.58	0.390	0.182	0.106	0.018
	(1693.19)	(1.771)	(30.70)	(106.36)	(548.77)	(0.191)	(0.175)	(0.148)	(0.063)
(12) 5	2563.23	7.286	70.14	213.31	1006.73	0.347	0.141	0.074	0.011
	(3589.15)	(1.506)	(32.81)	(129.82)	(793.74)	(0.171)	(0.146)	(0.118)	(0.049)

Numbers in parentheses are standard deviations.

The expanding technological diversity of more distant indirect collaborators of an inventor reflects the shrinking commonality in technological specialization between them and the inventor. To see this, let us compute the average Jaccar index between the technological specialization S_{it} of inventor i and those of his or her ℓ-th indirect collaborators $j \in N_{it}^{\ell}$ in period t:

$$j_{it}^{\ell} = \frac{1}{n_{it}^{\ell}} \sum_{j \in N_{it}^{\ell}} \frac{|S_{it} \cap S_{jt}|}{|S_{it} \cup S_{jt}|} \in [0,1] \tag{7.5}$$

where $n_{it}^{\ell} \equiv |N_{it}^{\ell}|$. A larger value of j_{it}^{ℓ} implies higher average similarity in technological specialization between inventor i and his or her ℓ-th indirect collaborators. In particular, it takes 0 if their specializations do not overlap (i.e., $S_{it} \cap S_{jt} = 0$ for all $j \in N_{it}^{\ell}$), while it takes 1 if they are identical (i.e., $S_{it} = S_{jt}$ for all $j \in N_{it}^{\ell}$).

Columns 6-9 of Table 7.1 show the average values of j_{it}^{ℓ} in terms of IPC sections, classes, subclasses and subgroups, respectively. These values between an inventor and his or her direct collaborators are on average 0.71, 0.57, 0.40 and 0.33 (0.76, 0.65, 0.59 and 0.43) in period 1 (period 2), respectively (rows 1 and 7 of column 6-9 in Table 7.1). The numbers of technological categories that an inventor shares with his or her collaborator are on average 1.39, 1.53, 1.62 and 2.05 (1.26, 1.35, 1.41 and 1.71) at IPC section, class, subclass and subgroup levels, respectively in period 1 (period 2).

Between an inventor and his or her fifth-indirect collaborators, however, the commonality of technological specialization is substantially smaller. The corresponding Jaccar indices reduce to 0.34, 0.13, 0.06 and 0.01 (0.35, 0.14, 0.07, 0.01), respectively in period 1 (period 2) (rows 6 and 12 of columns 6-9 in Table 7.1). The numbers of technological categories that an inventor shares with his or her fifth-indirect collaborator are as small as 0.79, 0.40, 0.24 and 0.07 (0.80, 0.43, 0.25 and 0.07) on average, respectively in period 1 (period 2).

Thus, we conclude that as long as inventors are sufficiently far apart on the collaborator network, say between the fifth-indirect collaborators, their research fields are virtually irrelevant. Note, in addition, that the firm- and location-specific effects underlying the similarity in productivity among indirect collaborators in the outcome of models (5.1) and (5.10) are controlled for by inventors' fixed effects as well as a variety of local factors. Hence, there should remain little concern about the endogeneity due to unobserved factors behind the productivity similarity among indirect collaborators.

7.3 Relevance

In this section, we argue that the relevance of our instruments comes essentially from the assortative matching by productivity at the firm level which is exogenous to the inventor collaboration in our data. In Section 7.3.1, we start by showing evidence for assortative matching among firms in their investment decisions by various financial performance indicators as well as worker productivity. We then show in Section 7.3.2 that the pool of potential collaborators for an inventor is largely confined within a single firm or its affiliated partners; hence, it can be considered as exogenous to each inventor, and its members have similar productivities.

7.3.1 Assortative matching of firms by worker productivity

If firms with investment relations as well as firms and their workers exhibit assortative matching by productivity, we expect that the productivities of inventors in these matched firms would be positively correlated. Evidence for assortative matching between firms and workers can be found in the existing literature (e.g., Mendes et al., 2010; Bartolucci and Devicienti, 2013; Dauth et al., 2016).[39] While direct evidence for assortative matching among firms is not available from the existing literature, we found supportive evidence from the financial and ownership data for Japanese firms (Tokyo Shoko Research, 2014).[40]

From 315,347 firms with financial information available in Japan in 2014, we identified 58,634 firm pairs with investment relationships. We then constructed an (undirected) network of firms with each firm as a node and each firm pair with investment relation as an edge. Table 7.2 shows average values of Spearman's rank correlations for average wage together with four financial indices between a firm and its directly/indirectly connected firms in the network of investment relations. The "indirectness" is defined analogously to that of the inventor network, so that value 0 indicates the direct investment relation, while value $j \geq 1$ indicates the j-th indirect investment relation.

Table 7.2: Rank correlations of financial indices between firms with ownership

Indirectness	Avg. wage	VA/worker	Capital-asset ratio	Pretax profit-asset ratio	Third-party evaluation
(1) 0	0.1267	0.0923	0.1824	0.1465	0.2577
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(2) 1	0.0930	0.0416	0.0490	0.0280	0.0926
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(3) 2	0.0260	0.0087	0.0045	0.0067	0.0132
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(4) 3	0.0121	-0.0017	0.0010	0.0019	-0.0086
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(5) 4	0.0005	-0.0038	-0.0023	0.0013	-0.0216
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(6) 5	-0.0017	-0.0026	-0.0013	0.0003	-0.0263
	(0.0000)	(0.0000)	(0.0000)	(0.0064)	(0.0000)

(*i*) The numbers in the parentheses are p-values of two-sided tests. (*ii*) "Avg. wage" represents the average nominal wage per (regular) worker; "VA/worker" the value added per worker; "Capital-asset ratio" the owned capital to total asset ratio; "Pretax profit-asset ratio" the pretax profit to total asset ratio; Third-party evaluation is the score ranging in [0,100] based on over 200 financial indices provided by the Tokyo Shoko Research.

One can see that the firms with investment relations exhibit positive correlations in the listed financial indices as well as average wage of workers (row 1). While the correlation quickly diminishes for more distant indirect partners, the relatively high correlations persist up to the first-indirect relation. For example, the correlations are 0.126 and 0.093 for average wage among firms with direct and first-indirect relations,

[39]See, e.g., Eeckhout and Kircher (2018) for a theoretical model.

[40]There are indirect evidences in the literature for assortative matching by productivity among firms. Namely, Bettencourt et al. (2007); Gaubert (2018); Dauth et al. (2016) have shown evidence for spatial sorting of firms and workers by productivity, while Nakajima et al. (2012) and Otazawa et al. (2018) have shown that firms with transaction linkages are geographically concentrated. See, for example, Mori and Turrini (2005); Behrens et al. (2014) for theoretical models of spatial sorting.

respectively.[41] It follows that firms with direct and first-indirect investment relations exhibit relatively stronger assortative matching by their financial performance as well as worker productivity. Since their workers include their inventors, the inventor productivities are expected to be correlated among these firms.

7.3.2 Collaboration and firm affiliation of inventors

If the size of a firm/establishment in period t is defined by the number of inventors who belong to the firm/establishment at some point in the period, the average and median of the firm size are 26,923 and 7,757 (23,025 and 8,207), while those of the establishment size are 3,500 and 1,059 (2,972 and 962) in period 1 (period 2).

In Table 7.3, columns 1 and 2 show the average shares of the ℓ-the indirect collaborators of an inventor in the same firm as the inventor ($\ell = 0, 1, \ldots, 5$) and columns 3 and 4 shows similar shares for establishments in periods 1 and 2, respectively. It is remarkable that on average more than 80% of collaborators are confined not only within a single firm but also within a single establishment. Although the shares steadily decrease as ℓ increases, they still remain as high as 25.6% and 31.7% for the fifth-indirect collaborators for firms, and 20.6% and 25.3% for establishments in periods 1 and 2, respectively.

Table 7.3: Firm affiliations of inventors

	(1)	(2)	(3)	(4)	(5)	(6)
	Same firm share		Same establishment share		Path length to firm	
Indirectness	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2
(1) 0	0.819	0.824	0.811	0.814	0.453	0.399
	(0.248)	(0.261)	(0.263)	(0.275)	(0.498)	(0.490)
(2) 1	0.721	0.731	0.687	0.694	0.887	0.789
	(0.271)	(0.289)	(0.303)	(0.319)	(0.568)	(0.626)
(3) 2	0.616	0.640	0.565	0.584	1.179	1.107
	(0.294)	(0.308)	(0.325)	(0.338)	(0.543)	(0.630)
(4) 3	0.501	0.539	0.440	0.473	1.410	1.354
	(0.302)	(0.316)	(0.328)	(0.340)	(0.534)	(0.625)
(5) 4	0.377	0.430	0.318	0.360	1.633	1.584
	(0.293)	(0.311)	(0.307)	(0.325)	(0.525)	(0.618)
(6) 5	0.256	0.317	0.206	0.253	1.843	1.794
	(0.260)	(0.290)	(0.260)	(0.291)	(0.499)	(0.602)

(*i*) Numbers in parentheses are standard deviations. (*ii*) "Same firm share" and "same establishment share" are the shares of ℓ-th indirect collaborators ($\ell = 0, 1, \ldots, 5$) of an inventor who belong to the same firm and the same establishment as the inventor, respectively. (*iii*) "Path length to firm" means the average number of firms on the shortest path from an inventor to the ℓ-th indirect collaborator on the research collaboration network of firms.

To see how many firms are involved to reach the ℓ-th indirect collaborators, we construct the collaboration network of firms with each firm as a node and each pair of firms conducting a joint patent development as an edge. Columns 5 and 6 of Table

[41] Note that since "workers" include all regular employees, the correlation in the average wage here tends to understate that in average wage among skilled workers including inventors.

7.3 list the values for the average shortest-path length from the firm of an inventor to the firm of his or her ℓ-th indirect collaborator on this collaboration network in periods 1 and 2, respectively. Although the shortest-path length, i.e., the smallest number of distinct firms involved, to reach the ℓ-th indirect collaborators, increases as ℓ increases in both periods, it still remains to be smaller than two even at the fifth-indirect collaborators. Provided that joint R&Ds take place more often among firms with closer investment relations, it implies that up to the fifth-indirect collaborators on average belong to firms with closer than the second-indirect investment relation.

It follows that the pool of potential collaborators for an inventor is confined mostly in a single firm or its closely affiliated firms.[42] This in turn indicates that the R&D is largely driven by firms, rather than inventors, and that the pool of potential collaborators is essentially exogenous to inventors in our regression, since we focus on the inventors who do not change their affiliation to firms and establishments.[43]

Since we already know that the firms with direct investment relations exhibit assortative matching in terms of worker productivity, it follows that in the exogenous pool of potential collaborators, the productivities of inventors in the affiliated firms are expected to be positively correlated, and hence the values of differentiated knowledge of collaborators, $\ln k_{it}^D$, of inventors in these affiliated firms are expected to be positively correlated.

As for Δn_{it}, recall Observation 2 in Section 2.3 that inventors with higher productivities conduct more active recombination of collaborators. As a result, the size of the collaborator recombination, Δn_{it}, is expected to be relatively similar among indirect collaborators with similar productivities. Yet, between inventor i and his or her indirect collaborator j, the relevance between Δn_{it} and Δn_{jt} induced by the assortative matching among firms and workers is weaker than that between k_{it}^D and k_{jt}^D, since the former pair are not related to the productivities of i and j directly unlike the latter pair. Thus, rather than (7.3), we adopt the alternative instrument given by (7.4) for Δn_{it} that puts more weights on the indirect collaborators who are more frequently connected to i.

8 Regression results

This section presents our main regression results for models (5.1), (5.9) and (5.10) under the quality- and novelty-adjusted productivity measures. In all the regressions

[42] More demanding matching between inventors and their firm/establishment affiliations in the Japanese patent applications by Inoue et al. (2016) indicates that more than 90% of the inventions take place within a single establishment, which further supports our argument.

[43] It is contrastive to academic research which is essentially driven by individual researchers, and the set of potential collaborators highly depends on the effort and ability of individual researchers, rather than their affiliated institutions.

conducted, the fixed effects of inventors, periods, and IPC classes (see Section 6.1.2) are controlled for. The local factors described in Section 6.3 except for residential population are constructed for a circle with a 1 km radius around each inventor to approximate establishment-specific effects, while it is set to 20 km for residential population to account for urban environment.

Standard errors in all the regressions are clustered by UAs (refer to Section 6.3),[44] since the productivities of collaborative activities within each UA are expected to be influenced by the stochastic shocks specific to the UA. In this context, since the instruments $\ln k_{it}^{D,\mathrm{IV}_\ell}$ for $\ln k_{it}^{D}$ in (5.1) and (5.9) as well as $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ for $\ln \Delta n_{it}$ in (5.10) involve inventors located in different UAs, one might suspect that standard cluster-robust standard errors are incorrect because the instruments for any inventor i might be correlated with errors ε_{jt} in (5.1), ε_{jt}^{m} in (5.6) and ϵ_{jt} in (5.10) for any inventor j even if inventors i and j are located different UAs. However, we consider that the standard cluster-robust standard errors still provide correct standard errors, since the inventor fixed effects are controlled in all the regressions encompass UA specific fixed effects, and that makes the errors free from the correlation with UAs, while allowing for standard errors to vary across UAs.

8.1 The Berliant-Fujita model

Table 8.1 summarizes the regression results for model (5.1), with columns 1-5 (6-10) presenting the results for quality-adjusted (novelty-adjusted) productivity. Columns 1 and 6 report the results from the ordinary least squares (OLS) regression for quality- and novelty-adjusted productivity, respectively, while the rest report those from the 2SLS-IV regressions. For the IV regressions, we used the third to fifth indirect collaborators to construct IVs for $\ln k_{it}^{D}$. More specifically, we used all three instruments, $\ln k_{it}^{\mathrm{IV}_\ell}$ for $\ell = 3,4$ and 5, in column 2 (column 7), while we used only one of $\ell = 3,4$ and 5 in columns 3, 4, and 5 (8, 9, and 10), respectively for quality- (novelty-) adjusted productivity.[45] To make the results comparable, the observations are restricted to the set of 58,464 inventors (rather than the 107,724 considered in Sections 2 and 6) with at least one fifth indirect collaborator.[46]

The OLS results confirm our earlier finding in Section 2 on the implication from

[44] As R&D activities are highly urban and agglomerative, almost all observations in I are found in UAs. In fact, among the 56,464 inventors in I with at least one fifth-indirect collaborators, and are chosen to be the basic set of observations in all the IV regressions in this section, only four inventors locate outside the UAs.

[45] Aside from the theoretical gap between (5.1) and (5.10) pointed out in Section 5.3, it in fact looks as if the instrument $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ for $\ln \Delta n_{it}$ also works as an instrument for $\ln k_{it}^{D}$ in the estimation of (5.1) because $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ has relevance with $\ln k_{it}^{D}$ via (5.10). However, the relevance turned out to be rather weak between $\Delta n_{it}^{\mathrm{IV}_\ell}$ and k_{it}^{D}, although Δn_{it} has positive significant effect on k_{it}^{D}.

[46] The basic properties of each variable remain the same, as described in Table 6.1.

Berliant and Fujita (2008) who predicted a positive effect of collaborators' differentiated knowledge, $\ln k_{it}^{D}$ (row 1, columns 1 and 6). The estimated positive effect of the knowledge stock, $\ln k_{it}$, of an inventor (row 2, columns 1 and 6) and the negative effect of its squared term, $(\ln k_{it})^2$ (row 3, columns 1 and 6), are consistent with the positive but decreasing returns of learning-by-doing from the extant technologies discussed in Sections 2 and 5.

However, since $\ln k_{it} > 0$ from the definition of k_{it} (≥ 1) in our data, the second-order effects appear to dominate the first-order effects; in other words, the net effect of the knowledge stock appears to be mostly negative. The overall negative effects associated with the knowledge stock imply that the positive learning-by-doing effects are dominated by the negative effects from imitations and obsolescence, which accounts for the persistent downward pressure on inventor productivity pointed out in Observation 1 in Section 2.2.

Table 8.1: Regression results for (5.1) (Dependent variable: $\ln y_{it}$)

	Quality					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k_{it}^{D}$	0.163***	0.286***	0.286***	0.287***	0.273***	0.164***	0.344***	0.341***	0.353***	0.377***
	(0.0102)	(0.0254)	(0.0257)	(0.0353)	(0.0399)	(0.00493)	(0.0310)	(0.0335)	(0.0296)	(0.0629)
(2) $\ln k_{it}$	0.110***	0.0931***	0.0931***	0.0929***	0.0949***	0.147***	0.114***	0.115***	0.113***	0.108***
	(0.0153)	(0.0119)	(0.0119)	(0.0143)	(0.0165)	(0.0172)	(0.0228)	(0.0229)	(0.0226)	(0.0248)
(3) $(\ln k_{it})^2$	-0.0890***	-0.0820***	-0.0820***	-0.0820***	-0.0828***	-0.195***	-0.178***	-0.178***	-0.177***	-0.175***
	(0.00967)	(0.00868)	(0.00865)	(0.00954)	(0.00991)	(0.00926)	(0.00594)	(0.00564)	(0.00665)	(0.0108)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.171***	0.117*	0.117*	0.117*	0.123**	0.310***	0.200**	0.202**	0.195**	0.180***
	(0.0579)	(0.0633)	(0.0635)	(0.0597)	(0.0540)	(0.0913)	(0.0939)	(0.0965)	(0.0887)	(0.0672)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0272***	0.0256***	0.0256***	0.0256***	0.0258***	0.0420***	0.0364***	0.0365***	0.0362***	0.0354***
	(0.00786)	(0.00679)	(0.00679)	(0.00664)	(0.00670)	(0.0156)	(0.0127)	(0.0128)	(0.0125)	(0.0120)
(6) $\ln a_{it}^{\mathrm{MNF}e}$	0.0149***	0.0240***	0.0240***	0.0240***	0.0230***	-0.00859	0.0132	0.0128	0.0143	0.0172
	(0.00566)	(0.00438)	(0.00436)	(0.00533)	(0.00598)	(0.0105)	(0.00989)	(0.00955)	(0.0108)	(0.0158)
(7) $\ln a_{it}^{\mathrm{MNF}o}$	0.00832	0.00522	0.00522	0.00520	0.00555	-0.00362	-0.00512	-0.00509	-0.00519	-0.00539
	(0.00581)	(0.00804)	(0.00806)	(0.00779)	(0.00732)	(0.00552)	(0.00721)	(0.00717)	(0.00732)	(0.00761)
(8) $\ln a_{it}^{\mathrm{POP}}$	-0.449	-0.660	-0.660	-0.661	-0.637	0.793*	0.0701	0.0837	0.0346	-0.0611
	(0.519)	(0.490)	(0.490)	(0.493)	(0.470)	(0.442)	(0.415)	(0.427)	(0.390)	(0.358)
(9) τ_1	0.227***	0.173***	0.173***	0.172***	0.178***	0.304***	0.173***	0.175***	0.166***	0.149***
	(0.0159)	(0.0150)	(0.0149)	(0.0213)	(0.0245)	(0.0307)	(0.0382)	(0.0403)	(0.0352)	(0.0477)
(10) R^2	0.151					0.184				
(11) Hansen J p-val.		0.928					0.768			
(12) 1st stage F		727.1	2178	1080	509.6		557.6	1590	918.7	471.4
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

We now turn to the IV results. For all the choices of IVs, the first-stage F values are large (row 12, columns 2-5 and 7-10), meaning that the relevance of the IVs does not seem to be weak (see Table D.1 in Appendix D for the results of the first-stage regressions). To confirm the exogeneity of the IVs, we used $\ln k_{it}^{\mathrm{IV}\ell}$ for all $\ell = 3, 4$ and 5 in columns 2 and 7 for quality- and novelty-adjusted productivities, respectively and conducted Hansen's (1982) J test for overidentifying restrictions. The p-values of the test are 0.928 and 0.768 for quality- and novelty-adjusted productivities, respectively (row 11, columns 2 and 7), meaning that the exogeneity of the IVs cannot be rejected.[47] Moreover, the estimated coefficients for the alternative choices of the IVs are remarkably similar (compare columns 2-5 with columns 7-10), which is also indicative of these IVs being reasonably exogenous.

The comparison between the OLS and IV results shows that the negative net effect of the knowledge stock persists in the IV result; hence, our explanation above for the OLS results continues to be valid.

For the effect of $\ln k_{it}^{D}$, we found downward bias in the OLS regression (compare columns 1 and 2-5 with columns 6 and 7-10 in row 1).[48] A possible explanation for the bias is that a more productive inventor attracts (or is assigned by his or her firm) a larger number of relatively unexperienced collaborators and thus tends to end up with more collaborators with lower productivity than he or she actively chose to work with. The removal of this reverse causality left a larger positive selection effect in the estimated coefficient of $\ln k_{it}^{D}$.

For the OLS regression, this selection effect may be partly picked up by the effect of the local concentration of inventors, $\ln a_{it}^{\mathrm{INV}}$, which has upward bias (compare columns 1 and 2 with columns 6 and 7 in row 4). Larger differentiated knowledge is not necessarily associated with a larger potential inventor population unless inventors actively choose to start new collaborations. However, a larger inventor concentration should naturally induce more fruitful collaborations, resulting in larger differentiated knowledge from collaborators, than a smaller one does. As a consequence, in the IV result, the part of the OLS estimate of the coefficient of $\ln a_{it}^{\mathrm{INV}}$ for which the collaborator recombination is responsible is absorbed into the coefficient of $\ln k_{it}^{D}$. What is left in the estimated effect of $\ln a_{it}^{\mathrm{INV}}$ may be interpreted as the positive spillover effect from the local inventor concentration. Specifically, a 10% increase in the inventor concentration results in 1.2% and 1.8-2.0% increases in quality- and novelty-adjusted productivity, respectively.

It is intuitive that the concentration of R&D expenditure has a persistent positive effect for all the specifications (row 5), where its 10% increase raises quality- and novelty-adjusted productivities by 0.26% and 0.35-0.36%, respectively, while the size of manufacturing output has essentially no impact on innovation productivity.

The positive significant effects of local manufacturing employment on quality-adjusted productivity (row 6, columns 1-5), where its 10% increase raises productivity by 0.23-0.24%, reflect the fact that innovations are linked to production; and citations are often made by the related production units of nearby firms. On the contrary, the manufacturing employment concentration is insignificant for novelty-adjusted productivity (row 5, columns 6-10), as technological novelty is not neces-

[47] Of course, this result of Hansen's J test by no means is sufficient to guarantee the exogeneity of the instruments, if all the instruments are subject to the same type and magnitude of bias.

[48] Akcigit et al. (2018) reported a similar downward bias on the effects of interaction levels on innovation productivity within a patent team.

sarily directly related to present production levels.

The local concentrations of residential population do not have a significant influence on inventor productivity as expected.

The estimated coefficients of $\ln k_{it}^{D}$ for the IV regression are persistently positive, 0.27-0.29 (0.34-0.38) for quality-adjusted (novelty-adjusted) productivity, but below 1 (row 1, columns 2-5 and 8-10), which is consistent with the Berliant-Fujita model. This finding indicates decreasing returns to the differentiated knowledge of collaborators, as the benefit from collaborators' differentiated knowledge will eventually be dominated by that of common knowledge with collaborators as well as the differentiated knowledge of the inventor him- or herself.

8.2 Quality/novelty and quantity decomposition

In this section, the effect of each explanatory variable in (5.1) is decomposed into the fraction that accrues to the quantity and to the quality/novelty of his or her output. The result for the former is relegated to Appendix E.

The regression results are summarized in Table 8.2 which is organized similarly to Table 8.1, except for the dependent variable. The first-stage of the regression is shared with (5.1). To confirm the exogeneity of the IVs, we used $\ln k_{it}^{\mathrm{IV}_\ell}$ for all $\ell = 3, 4$ and 5 in columns 2 and 7 for quality- and novelty-adjusted productivities, respectively and conducted Hansen's (1982) J test for overidentifying restrictions. The p-values of the test are 0.419 and 0.314 for quality- and novelty-adjusted productivities, respectively (row 11, columns 2 and 7), meaning that the exogeneity of the IVs cannot be rejected.

Together with the results summarized in Table 8.1, the results from the present regressions in Table 8.2 reveal the extent to which each explanatory variable contributes to quality/novelty and to the quantity in collaborative knowledge creation.

For the differentiated knowledge of collaborators, whereas we find that its contribution is mostly (more than 90%) attributed to increasing the quantity, rather than the quality, of research output under the quality-adjusted productivity measure (compare row 1 and columns 2-5 in Tables 8.1 and 8.2), as large as around 65% of the contribution accrues to increasing the novelty, rather than the quantity, of research output under the novelty-adjusted productivity measure (compare row 1 and columns 7-10 in Tables 8.1 and 8.2).

This result indicates that the collaborators' differentiated knowledge is an especially effective source of technological novelty, and thus, appears to be the key factor for inducing the technological shift of an inventor to a new niche, which is consistent with Berliant and Fujita (2008) as well as Horii (2012).

Table 8.2: Regression results for (5.1) (Dependent variable: $\ln y^q_{it}$)

	Quality					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k^D_{it}$	0.0273***	0.0264**	0.0269**	0.0154	0.00321	0.119***	0.230***	0.231***	0.221***	0.247***
	(0.00169)	(0.0120)	(0.0119)	(0.0192)	(0.0221)	(0.00278)	(0.0143)	(0.0146)	(0.0168)	(0.0330)
(2) $\ln k_{it}$	0.0104*	0.0106	0.0105	0.0121	0.0138*	0.0364	0.0162	0.0161	0.0179	0.0132
	(0.00578)	(0.00714)	(0.00710)	(0.00832)	(0.00758)	(0.0240)	(0.0274)	(0.0274)	(0.0272)	(0.0272)
(3) $(\ln k_{it})^2$	-0.00549***	-0.00554***	-0.00551***	-0.00616***	-0.00684***	-0.108***	-0.0976***	-0.0976***	-0.0985***	-0.0961***
	(0.00101)	(0.00138)	(0.00136)	(0.00189)	(0.00202)	(0.00520)	(0.00651)	(0.00655)	(0.00653)	(0.00657)
(4) $\ln a^{\mathrm{INV}}_{it}$	-0.0364***	-0.0361**	-0.0363**	-0.0313**	-0.0260**	0.0719**	0.00411	0.00368	0.00967	-0.00596
	(0.0126)	(0.0141)	(0.0142)	(0.0137)	(0.0102)	(0.0318)	(0.0392)	(0.0397)	(0.0383)	(0.0400)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	-0.00252	-0.00251	-0.00252	-0.00236	-0.00220	0.0118*	0.00839	0.00837	0.00868	0.00789
	(0.00414)	(0.00421)	(0.00422)	(0.00409)	(0.00377)	(0.00638)	(0.00570)	(0.00569)	(0.00569)	(0.00598)
(6) $\ln a^{\mathrm{MNF}e}_{it}$	0.0271***	0.0271***	0.0271***	0.0262***	0.0254***	0.00798	0.0215**	0.0216**	0.0204**	0.0235**
	(0.00559)	(0.00521)	(0.00520)	(0.00531)	(0.00641)	(0.00850)	(0.00929)	(0.00919)	(0.00949)	(0.0112)
(7) $\ln a^{\mathrm{MNF}o}_{it}$	0.00861	0.00863	0.00862	0.00891*	0.00921*	-0.00636	-0.00728	-0.00729	-0.00720	-0.00742
	(0.00556)	(0.00534)	(0.00535)	(0.00518)	(0.00526)	(0.00458)	(0.00656)	(0.00657)	(0.00637)	(0.00684)
(8) $\ln a^{\mathrm{POP}}_{it}$	-0.582**	-0.580**	-0.581**	-0.562**	-0.541**	0.610	0.163	0.160	0.200	0.0971
	(0.238)	(0.240)	(0.239)	(0.250)	(0.251)	(0.440)	(0.477)	(0.476)	(0.479)	(0.497)
(9) τ_1	0.101***	0.102***	0.101***	0.107***	0.112***	0.151***	0.0699***	0.0694***	0.0765***	0.0578*
	(0.0180)	(0.0223)	(0.0223)	(0.0245)	(0.0200)	(0.0202)	(0.0210)	(0.0209)	(0.0225)	(0.0323)
(10) R^2	0.086					0.140				
(11) Hansen J p-val.		0.419					0.314			
(12) 1st stage F		727.1	2178	1080	509.6		557.6	1590	918.7	471.4
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

The decompositions of the effects of other explanatory variables are also worth explanations. For both quality and novelty-adjusted productivity measures, the inventor as well as R&D expenditure concentrations exhibit positive significant effect on the quantity but not on the quality of inventions (rows 4 and 5 in Tables 8.2 and E.1). The effects of manufacturing employment and production concentrations are also similar between quality- and novelty-adjusted cases. But, they tend to raise the quality rather than the quantity of inventions (rows 6 and 7 in Tables 8.2 and E.1). The former result suggests that positive externalities from researcher agglomeration primarily promote starting inventions, whereas the latter result may reflect the tendency that the proximity to the manufacturing concentration and production promotes more targeted inventions with higher quality and novelty.

The results of our regressions so far identified the causal relation suggested by the Berliant-Fujita model behind the correlation between collaborators' differentiated knowledge and the productivity of inventors in Observation 2 in Section 2.3, except for the linkage between the collaborator recombination of an inventor and the amount of differentiated knowledge of his or her collaborators that will be established in Section 8.3.[49] It follows that technological shift, Δs_{it}, which was found to be correlated with higher productivity in Observation 2, is in fact intentionally directed toward less explored niches because of inventors' (or firms') quest for more novel invented technologies. The technological shift caused by utilizing collaborators'

[49] Although we used inventor productivity, $\bar{y}$, in Section 2 rather than pairwise productivity, y_{it}, these are highly correlated, with correlation coefficients of 0.73 and 0.76 in periods 1 and 2, respectively. Thus, the observations made for $\bar{y}$ in Section 2 basically apply to y_{it} as well.

differentiated knowledge appears to be a means to overcome the negative effects of the past knowledge stock of inventors pointed out in Observation 1. This part of the result is missing from the Berliant-Fujita model in which all knowledge is assumed to be symmetric. However, this finding agrees with the theoretical result of Horii (2012), who considered a more realistic economy with demand for new technologies.

8.3 Recombinations and differentiated knowledge of collaborators

This section presents the results for model (5.10), which incorporated the fundamental causality assumed in the Berliant-Fujita model that the collaborator recombination is an effective means to collect novel ideas for innovations. Table 8.3 summarizes the regression results. This table is organized similarly to Table 8.1, except that the dependent variable is $\ln \Delta k_{it}^{D}$, and $\ln \Delta n_{it}^{\mathrm{IV}\ell}$ for $\ell = 3,4$ and 5 serve as the IVs for an endogenous variable, $\ln \Delta n_{it}$.

The OLS estimates suggest a positive effect of collaborator recombination on the size of collaborators' differentiated knowledge (row 1, columns 1 and 6) as expected. However, given the correlations among inventor productivity, collaborators' differentiated knowledge, and collaborator recombinations underlying the innovations, the OLS estimates may be severely biased due to endogeneity. The IV estimates in columns 2-5 and 7-10 indicate that this is indeed the case.

Now, we look at the IV results in detail. For all the different choices of IVs, the first-stage F values are large (row 12, columns 2-5 and 7-10), suggesting that the relevance of the IVs is not weak (see Table D.2 in Appendix D for the results from the first-stage regressions). To confirm the exogeneity of the IVs, we used $\ln \Delta n_{it}^{\mathrm{IV}\ell}$ for all $\ell = 3,4$ and 5 in columns 2 and 7 for quality- and novelty-adjusted productivities, respectively and conducted Hansen's (1982) J test for overidentifying restrictions. The p-values of the test are 0.255 and 0.363 for quality- and novelty-adjusted productivities, respectively (row 11, columns 2 and 7), meaning that the exogeneity of the IVs cannot be rejected.[50] The estimated coefficients for the alternative choices of the IVs are less stable than those for model (5.1), but they agree with each other qualitatively (compare columns 2-5 with columns 7-10).

There is substantial downward bias in the coefficient estimate for $\ln \Delta n_{it}$ from the OLS (compare columns 1 and 2 with columns 6 and 7 in row 1). For the OLS result, a part of the effect of collaborator recombination appears in that of local inventor concentration, since a larger inventor concentration implies a larger pool of potential collaborators. The downsized effect of $\ln a_{it}^{\mathrm{IV}}$ in the IV regression is consistent with this interpretation (compare columns 1 and 2-5 with columns 6 and 7-10 in row 5).

Another source of the bias is reverse causality. A higher productivity for an

[50] The same caveat stated in footnote 47 applies here.

inventor is on average associated with the larger differentiated knowledge of his or her collaborators as well as a larger stock of knowledge. This bias appears to be reflected in the estimated coefficient of the knowledge stock, $\ln k_{it}$, which has substantial upward bias in the OLS (compare columns 1 and 2-5 with columns 6 and 7-10 in row 2). Once the endogeneity of $\ln \Delta n_{it}$ is controlled for, we find that the first-order effect of the knowledge stock almost disappears (columns 2-5 and 7-10 of row 2), and instead the second-order effect becomes positive significant (columns 2-5 and 7-10 of row 3); thus, the effect of the knowledge stock exhibits increasing returns. The size of differentiated knowledge is then not necessarily associated with a larger number of new collaborators.

On the one hand, a highly established inventor with a large knowledge stock can attract highly able collaborators selectively even without a large replacement of collaborators. On the other hand, an inventor with only a small stock of knowledge should place a large effort to find relevant collaborators for successful inventions (or his or her firm should arrange so), which in turn results in a large number of new collaborators. Other local factors play relatively minor roles.

We find that the elasticities of average quality- and novelty-adjusted differentiated knowledge of collaborators with respect to the recombination of collaborators for an inventor are around 1.4 and 1.8, respectively. While these estimated elasticities are greater than 1, since the pairwise research productivity exhibits decreasing returns in the input of collaborators' differentiated knowledge, the positive effect of the collaborator recombination on inventor productivity will be diminishing. More specifically, putting the results from (5.1) and (5.10) together, we found that a 10% increase in collaborator recombination induces 12-15% and 17-20% increases in quality- and novelty-adjusted differentiated knowledge of collaborators, which in turn results in the 3-4% and 6-8% increases in quality- and novelty-adjusted pairwise output of an inventor.

Taken together, we confirmed that collaborator recombinations are an effective means to acquire differentiated knowledge from new collaborators to facilitate invention, thereby identifying the causal relationship behind Observation 2. Moreover, the results of our regressions also accounted for the mechanism behind Observation 3 in Section 2.4. In other words, we found that the knowledge stock and collaborator recombination remain two effective means for an inventor to improve his or her productivity via collecting differentiated knowledge, even after controlling for the individual fixed effects. Inventors with a larger past achievement attract highly able collaborators with their large knowledge stocks, and thus can collect differentiated knowledge without large replacements of collaborators. Meanwhile, less experienced ones depend on relatively large recombination of collaborators to collect differentiated knowledge.

Taken all together, the rather intricate mechanism underlying the churning of inventor productivities in Observation 1 has been disentangled, and explained from the micro-level behavior of individual inventors.

Table 8.3: Regression results for (5.10) (Dependent variable: $\ln k_{it}^{D}$)

	Quality					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln \Delta n_{it}$	0.104***	1.372***	1.400***	1.523***	1.236***	0.244***	1.722***	1.718***	1.962***	1.714***
	(0.00626)	(0.0629)	(0.0748)	(0.119)	(0.132)	(0.00792)	(0.0847)	(0.0914)	(0.157)	(0.131)
(2) $\ln k_{it}$	0.131***	-0.0220	-0.0253	-0.0401	-0.00554	0.147***	-0.0313	-0.0308	-0.0601	-0.0303
	(0.0427)	(0.0669)	(0.0653)	(0.0814)	(0.0732)	(0.0338)	(0.0638)	(0.0632)	(0.0786)	(0.0669)
(3) $(\ln k_{it})^2$	-0.0364**	0.223***	0.229***	0.254***	0.195***	-0.0422***	0.261***	0.260***	0.310***	0.259***
	(0.0156)	(0.0197)	(0.0167)	(0.0422)	(0.0392)	(0.0161)	(0.0233)	(0.0223)	(0.0476)	(0.0360)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.387***	0.0138	0.00580	-0.0304	0.0539	0.515***	0.0800	0.0813	0.00957	0.0825
	(0.0916)	(0.0426)	(0.0467)	(0.0461)	(0.0507)	(0.118)	(0.103)	(0.107)	(0.0878)	(0.0939)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0134	0.000705	0.000432	-0.000799	0.00207	0.0320*	0.0172*	0.0172*	0.0148	0.0173*
	(0.0111)	(0.00478)	(0.00487)	(0.00571)	(0.00447)	(0.0165)	(0.00896)	(0.00895)	(0.00939)	(0.00915)
(6) $\ln a_{it}^{\mathrm{MNF}e}$	-0.0706***	-0.0139	-0.0127	-0.00720	-0.0200	-0.110***	-0.0436**	-0.0438**	-0.0329	-0.0440*
	(0.0220)	(0.0147)	(0.0151)	(0.0183)	(0.0137)	(0.0186)	(0.0219)	(0.0213)	(0.0302)	(0.0239)
(7) $\ln a_{it}^{\mathrm{MNF}o}$	0.0214	0.00814	0.00786	0.00657	0.00957	0.00221	-0.0133	-0.0132	-0.0158	-0.0132
	(0.0215)	(0.00992)	(0.0101)	(0.0110)	(0.00960)	(0.0265)	(0.00922)	(0.00923)	(0.0107)	(0.00920)
(8) $\ln a_{it}^{\mathrm{POP}}$	1.371	-0.552	-0.594	-0.780	-0.345	3.574***	1.332	1.338	0.968	1.345
	(1.043)	(1.229)	(1.217)	(1.403)	(1.273)	(1.137)	(1.050)	(1.037)	(1.247)	(1.103)
(9) τ_1	0.415***	0.514***	0.517***	0.526***	0.504***	0.698***	0.814***	0.814***	0.833***	0.814***
	(0.0269)	(0.0504)	(0.0525)	(0.0481)	(0.0428)	(0.0373)	(0.0285)	(0.0287)	(0.0310)	(0.0291)
(10) R^2	0.160					0.178				
(11) Hansen J p-val.		0.255					0.363			
(12) 1st stage F		237.7	639.9	338.5	253.9		237.7	639.9	338.5	253.9
(13) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

9 Robustness

In this section, we check robustness of our baseline results with emphasis on two aspects: the influence of firm-specific factors in Section 9.1 and the exogeneity condition for our baseline IVs in Section 9.2. We also investigate the sensitivity of our baseline results under the alternative definitions for inventor productivities as well as under the alternative spatial sizes at which local factors are defined in Section 9.3.

9.1 Influence of firm-specific factors

In this section, we examine the influence of the three time-varying properties of the firm to which each inventor belongs to. Let F_{it} be the set of inventors who belong to the same firm as inventor i at some point in period t, and let $F_{-i,t} \equiv F_{it} \setminus (N_{it} \cup \{i\})$, i.e., *$F_{it}$ excluding i and his or her collaborators.*

The first property taken into account is the *firm size*, $f_{it} = |F_{-i,t}|$, representing the magnitude of the R&D activities within the firm that inventor i belongs to, but outside the projects that an inventor and his or her collaborators are directly involved. Given that more than 80% of collaborations take place within a firm, the variation in k_{it}^{D} as well as that of Δn_{it} may simply reflect that of firm size in period t.[51]

The second property is the *technological scope* of a firm of a given inventor i defined by $s^f_{it} = |\cup_{j\in F_{it}} S_{jt} \backslash (\cup_{u\in N_{it}\cup\{i\}} S_{ut})|$ which counts the number of distinct technological categories in which patents are developed in the firm of inventor i excluding those associated with the patents developed by inventor i and by his or her collaborators. The values of f_{it} and s^f_{it} reflect the potential scale effect of a firm. For example, the availability of common research facilities, funding and other resource is the source of increasing returns, and there may also be interdisciplinary spillovers.

Table 9.1: Regression results with firm-specific factors

	(5.1) Dependent variable : $\ln y_{it}$				(5.10) Dependent variable : $\ln k^D_{it}$			
	Quality		Novelty		Quality		Novelty	
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5	(7)OLS	(8) IV3-5
(1) $\ln k^D_{it}$	0.155***	0.224***	0.156***	0.284***				
	(0.00932)	(0.0256)	(0.00517)	(0.0295)				
(2) $\ln \Delta n_{it}$					0.102***	1.425***	0.229***	1.725***
					(0.00582)	(0.0848)	(0.00636)	(0.126)
(3) $\ln f_{it}$	-0.0388	-0.0411	-0.202	-0.0858	0.00282	0.0385	-0.807***	-0.0370
	(0.0415)	(0.0340)	(0.124)	(0.120)	(0.109)	(0.0581)	(0.0681)	(0.0809)
(4) $\ln s^f_{it}$	0.0622***	0.0468**	0.233***	0.0883	0.230***	0.0257	1.006***	0.106
	(0.0157)	(0.0189)	(0.0450)	(0.0602)	(0.0325)	(0.0266)	(0.119)	(0.0880)
(5) $\ln y^f_{it}$	0.498***	0.487***	0.104**	0.186***	0.181*	0.647***	-0.493***	0.407**
	(0.0764)	(0.0693)	(0.0506)	(0.0522)	(0.108)	(0.144)	(0.158)	(0.160)
(6) $\ln k_{it}$	0.105***	0.0964***	0.133***	0.115***	0.115***	-0.0350	0.110***	-0.0468
	(0.0118)	(0.0104)	(0.0168)	(0.0211)	(0.0424)	(0.0580)	(0.0301)	(0.0517)
(7) $(\ln k_{it})^2$	-0.0887***	-0.0851***	-0.192***	-0.182***	-0.0332**	0.235***	-0.0348**	0.264***
	(0.00824)	(0.00813)	(0.00880)	(0.00578)	(0.0150)	(0.0152)	(0.0157)	(0.0205)
(8) $\ln a^{\mathrm{INV}}_{it}$	0.168***	0.145***	0.291***	0.234***	0.292**	-0.0255	0.350**	0.0287
	(0.0284)	(0.0348)	(0.0468)	(0.0559)	(0.137)	(0.0351)	(0.140)	(0.0559)
(9) $\ln a^{\mathrm{R\&D}}_{it}$	0.00949*	0.00997**	0.0268***	0.0259***	-0.00811	-0.0226***	0.00675	-0.00905
	(0.00517)	(0.00482)	(0.0101)	(0.00953)	(0.00707)	(0.00836)	(0.0125)	(0.0115)
(10) $\ln a^{\mathrm{MNF}_e}_{it}$	-0.0154***	-0.00920**	-0.0239***	-0.00710	-0.0880***	-0.0416*	-0.121***	-0.0589***
	(0.00449)	(0.00457)	(0.00829)	(0.00724)	(0.0220)	(0.0222)	(0.0209)	(0.0166)
(11) $\ln a^{\mathrm{MNF}_o}_{it}$	0.00436	0.00268	0.000106	-0.00348	0.0207	0.00151	0.0204	-0.00978
	(0.00405)	(0.00534)	(0.00715)	(0.00779)	(0.0261)	(0.0125)	(0.0310)	(0.00975)
(12) $\ln a^{\mathrm{POP}}_{it}$	-0.233	-0.298	0.170	-0.0790	0.664	-0.331	1.724	0.469
	(0.309)	(0.299)	(0.435)	(0.440)	(1.379)	(1.086)	(1.496)	(1.124)
(13) τ_1	0.129***	0.110***	0.162***	0.109***	0.248***	0.411***	0.394***	0.614***
	(0.0184)	(0.0157)	(0.0318)	(0.0394)	(0.0443)	(0.0354)	(0.0330)	(0.0277)
(14) R^2	0.158		0.187		0.169		0.192	
(15) Hansen J p-val.		0.957		0.813		0.309		0.388
(16) 1st stage F		621.2		396.7		203.4		174.2
(17) #Obs.	114,258	114,258	114,258	114,258	92,552	92,552	92,552	92,552

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

The third and the final property to be considered is the average pairwise productivity in the firm given by $y^f_{it} = (1/f_{it})\sum_{j\in F_{it}\backslash(N_{it}\cup\{i\})} y_{it}$ which may be correlated with the positive spillovers that prevail within a firm.

[51]Note that the firm size here is rather special, as it aggregates all the inventors affiliated with a given firm at some point in the given period. The firm size may be slightly overstated, since inventors who simply changed establishments within a firm in the same period counted multiple times. Nonetheless, it should reflect the basic variation in the number of inventors involved in a given firm.

Columns 1-4 and 5-8 in Table 9.1 show the regression results under (5.1) and those under (5.10), respectively with these additional controls on the RHS. For quality- and novelty-adjusted productivity measures under each model, the table shows the OLS and the IV results, where the IV results are shown only for the case in which IVs are constructed by using all the third-, forth- and fifth-indirect collaborators, since similar results are obtained when only one of them is used.

In all the specifications, the first-stage F values are reasonably large, so that the relevance appears to be strong as in the baseline case. In terms of the Hansen (1982)'s J-test, there is no evidence against the exogeneity of the instruments.

We found that the the coefficient of $\ln k_{it}^{D}$ in (5.1) is significantly different between the baseline and the current specifications (in both the OLS and IV results and under both the quality- and novelty-adjusted productivity), while the coefficient of $\ln \Delta n_{it}$ in (5.10) is not significantly different, except for the IV result for the quality-adjusted productivity.[52]

Nevertheless, the signs and significance levels of the explanatory variables in the baseline specifications do not change for both the OLS and IV results. Although both the productivity of each inventor and the knowledge level of his or her collaborators are positively influenced by the average inventor productivity in a firm, they still substantially depend on the individuals' decision on collaborations. A 10% increase in the differentiated knowledge of collaborators raises the inventor productivity by 2.24% and 2.84%, and a 10% increase in the collaborator recombination raises the amount of differentiated knowledge of collaborators by 14.3% and 17.3% under quality- and novelty-adjusted productivity, respectively, even after controlling for the average productivity of a firm. The results in this section thus underscore the relevance of the Berliant-Fujita model.

Regarding the effects of the firm specific factors, the IV results indicate that the effect of firm size is insignificant for both (5.1) and (5.10). While the IV estimate of the effect of technological diversity, s_{it}^{f}, of firms is positive significant for (5.1) under the quality-adjusted productivity (row 4 of column 2 in Table 9.1), it is insignificant otherwise. Hence, it is unlikely that the variation in the size or the scope of research of a firm is driving that in the amount of differentiated knowledge of collaborators and the magnitude of collaborator recombination.

The effect of the average pairwise productivity, $\ln y_{it}^{f}$, is positive significant under all the specifications. A 10% increase in the average pairwise productivity in a firm is associated with 4.9% and 1.9% increase in the quality- and novelty-adjusted average pairwise productivity of an inventor, respectively (row 5 of columns 2 and 4), while it is associated with 6.5% and 4.1% increase, respectively in the differentiated

[52]It is based on the Wald test in the GMM estimation which simultaneously estimates the baseline and current models with the 2SLS weighting matrix.

knowledge of collaborators (row 5 of columns 6 and 8).

Given that most collaborations take place within a firm, the former result suggests the presence of positive externalities from more productive colleagues within a firm and/or increasing returns in quality/novelty-adjusted differentiated knowledge of collaborators. The latter result, on the other hand, indicates that the given the same number of new collaborators, the differentiated knowledge of collaborators is larger in a more productive firm.

9.2 IVs based on indirect collaborators in different firms

Next, we consider alternative IVs for $\ln k_{it}^{D}$ and $\ln \Delta n_{it}$ which are the same as $\ln k_{it}^{D,\mathrm{IV}_\ell}$ in (7.3) and $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ in (7.4) except that the inventors in the same firm as i are excluded from N_{it}^{ℓ}. By using these IVs, we are supposed to be able to mitigate unobserved firm-specific factors associated with the IVs which may correlate with the error term.

The regression results for (5.1) and and (5.10) are shown in Tables F.1 and F.2, respectively in Appendix F.[53] The results remain qualitatively the same as those from the baseline analyses shown in Tables 8.1 and 8.3, except that the estimated coefficient for $\ln k_{it}^{D}$ in (5.1) is insignificant under the IVs constructed from the forth- and fifth-indirect collaborators for the case of quality-adjusted productivity measure (columns 4 and 5 in row 1 of Table F.1). This is not surprising provided that the correlations of worker productivity among firms attenuate rather quickly as firms become far from each other on the investment network as discussed in Section 7.3. It is thus natural to loose the relevance of the instrument for an endogenous variable for inventor i constructed from his or her indirect collaborators outside than inside the firm that he or she belongs to.

Nevertheless, weak IVs are not found under the novelty-adjusted productivity measures (columns 6-10 of Table F.1). Moreover, for all cases, if we use both baseline and present IVs constructed from the ℓ-th indirect collaborators for $\ell = 3, 4$ and 5, the null hypothesis of the Hansen (1982)'s J-test was not rejected, which also suggest that unobserved firm-specific factors are of minor concern.

For (5.10), we have qualitatively the same results under the alternative IVs as those obtained in the baseline analyses in Section 8.3 (refer to Tables 8.3 and F.2). Thus, our baseline results for (5.10) are found to be robust.

[53]The choices of IVs in Tables F.1 and F.2 are the similar to those in Tables 8.1 and 8.3, respectively.

9.3 Other robustness analyses

9.3.1 Alternative productivity measures

The regressions for (5.1) and (5.10) are conducted under four alternative measures of inventor productivity, where the output g_j of patent j in (2.1) is given by (*i*) the cited count in five years from publication, (*ii*) technological novelty based on the IPC subclass, (*iii*) count of patent claims;[54] or (*iv*) count of patents, i.e., $g_j = 1$ for all j. The regression results are relegated to Appendix G, where Tables G.2 and G.3 present the results from the second-stage regressions for models (5.1) and (5.10), respectively.

Under all alternative productivity measures, the signs and the magnitude of the estimated estimated coefficients for the explanatory variables are highly consistent with our baseline results in Tables 8.1 and 8.3. Hence, we conclude that our results are robust under the alternative measures of inventor productivity.

9.3.2 Differentiated knowledge of collaborators by IPC

Here, we consider the differentiated knowledge of collaborators defined by the IPC subgroups, instead of the productivity-based measures:

$$k_{it}^{D} = \frac{1}{n_{it}} \sum_{j \in N_{it}} \left| S_{jt} \backslash S_{it} \right| , \tag{9.1}$$

although it corresponds less precisely to the knowledge creation function (4.2). Regression results for (5.1) and (5.10) are relegated to Table H.1 in Appendix H.

In all the specifications, the signs and significance of the estimated coefficient values of the explanatory variables are highly consistent with our baseline results shown in Tables 8.1 and 8.3 for (5.1) and (5.10), respectively. Hence, we conclude that our results are robust even if the differentiated knowledge of collaborators are defined in terms of the IPC categories.[55]

9.3.3 Alternative radius values for local concentrations

Finally, we consider alternative radius values (5, 10 and 20km) to quantify the magnitude of local concentration of inventors, R&D and manufacturing and population around each individual inventors in (5.1) and (5.10). The regression results are relegated to Tables I.1-I.4 in Appendix I.

[54]Each claim indicates an aspect of the patent to be protected. Thus, its count reflects the technological novelty *within a patent*. While the claims are made by applicants, this is not an entirely subjective measure of quality since each claim incurs monetary costs.

[55]The qualitative results remain the same when the IPC subclass instead of subgroup is adopted to define the differentiated knowledge of collaborators as well as the knowledge stock of inventors.

The estimated coefficients of the key variables, $\ln k_{it}^{D}$ and $\ln \Delta n_{it}$ are robust under all the alternative radius values at which the size of local concentrations are measured. The direct effects of the local concentrations on the outcome variables are also generally robust under these alternative choice of radius values.

In particular, for (5.1), the effects of local concentrations are spatially confined for the inventor, R&D and manufacturing concentrations in the sense that the effect is significant up to 5km radius (if significant at all), while the effect of residential population concentration, $\ln a_{it}^{\mathrm{POP}}$, is insignificant for all the radius values for 5-20km. For (5.10), the negative effects of $\ln a_{it}^{\mathrm{INV}}$ and $\ln a_{it}^{\mathrm{MNF}_e}$ persist for larger radius values, possibly reflecting the tougher competition for human resources with co-localizing R&D activities as well as the manufacturing sector.

10 Discussion and further research directions

In this section, we summarize our findings and their implications, and discuss possible immediate extensions and further research directions.

10.1 The Berliant-Fujita mechanism and beyond

We have shown evidence consistent with the polyadic collaborative knowledge creation mechanism proposed by Berliant and Fujita (2008). To our knowledge, our work is the first to provide micro-econometric evidence for knowledge creation at the individual inventor level taking into account the endogeneity of collaborations.

We have also addressed two major counterfactual aspects of the Berliant-Fujita model, guided by Horii's (2012) result. One is that each inventor in their model belongs to a fixed network component in a typical steady state, meaning that polyadic interactions happen only within a given set of collaborators. However, in the data, the set of collaborators evolves for each agent over time, and the inter-temporal recombination of collaborators is found to revise inventors' technological expertise by meeting new agents and adopting their differentiated knowledge.

The other is that inventors in their model face no imitation or obsolescence of their technological knowledge since the number of potential knowledge is infinite and they are symmetric. In reality, however, we found negative significant effects from the knowledge stock of inventors on their productivity. If inventors stick to their past achievement, they most likely lose their present level of creativity in the long run. If, instead, agents are willing to explore new research directions by meeting new collaborators with different backgrounds from theirs, they are more likely to keep their creativity by shifting their technological expertise to unexplored niches. We have explained this realistic causal relationship by estimating the second and the third regression models, (5.9) and (5.10), in addition to the original Berliant-Fujita model (5.1). Specifically, collaborator recombinations are found to be effective in raising the quality as well as novelty of the collaborators' differentiated knowledge, thereby enhance the quality and novelty of research output of an inventor.

These additional results reveal a so far overlooked aspect of collaborative knowledge creation. Namely, the active collaborator recombinations are an effective strategy for a fledgling inventor to improve his or her research productivity, as well as for an established inventor to maintain his or her high productivity (although the latter can also utilize his or her large stock of knowledge).

This evidence has an important policy implication: Firms, cities, regions and countries that promote encounters and collaborations among individual inventors across organizations and institutions, despite the possibility of imitations and undesired diffusions, may have better chances to foster innovation there. While lower organizational and institutional barriers for research collaboration are not incompatible with the protection of intellectual property by patents, our finding supports more active coordination than divisions among researchers to encourage innovations.[56]

10.2 Extensions

Among a number of short-run and long-run extensions, we touch on three. First, it is an obvious interest to further investigate the role of firms and establishments in R&D activities. Since the financial resources for R&D are typically provided by firms, firm-specific patterns of collaborations and R&D policies could affect the productivity of individual inventors.[57] By matching the addresses of establishments in the patent database with those of the Census of Manufacturers, it is in principle possible to investigate the impact of patent development on firm productivity.

Second, the non-technological diversity among collaborators in terms of, for example, gender, age and cultural background may affect productivity. For example, Østergaard et al. (2011) and Inui et al. (2014) found positive influence of gender diversity in innovation productivity of Danish and Japanese firms, respectively.

Finally, it is intriguing to explore the differences in the location patterns of R&D activities and industries. It is argued that disproportionately large cities are typically associated with a concentration of knowledge-intensive activities (e.g., Davis and Dingel, 2017, 2018). However, the fundamental distinction between knowledge-intensive and non-intensive activities has not been made clear thus far.

From our findings, obviously knowledge-intensive R&D activities are expected to

[56]See Boldrin and Levine (2013) for a related survey of the literature arguing that the patent system hinders rather than promotes innovations.

[57]See Akcigit and Kerr (2018) for an initial attempt in this direction, as they distinguish R&D that is internal and external to firms and study the firm dynamics that arise from this distinction.

be more concentrated geographically given their incentive for frequent collaborator recombinations than industrial activities whose concentrations are typically induced by input-output linkages, demand, and production externalities.

Figure 10.1(a) plots the aggregate novelty-adjusted patent output and manufacturing output against the population size of a UA in period 1, where all values are expressed by shares in all UAs.[58] The solid and dashed lines indicate the fitted OLS lines for the patent count and manufacturing output plots, respectively. While both plots are super-linear in UA size (i.e., per-capita productivity is increasing in agglomeration size), it is substantially more so for patent output. In fact, doubling the population size of a UA raises R&D productivity by 2.5 times,[59] while raising manufacturing productivity only by 1.2 times.

Figure 10.1(b) plots the diversity in IPC subclasses of applied patents as well as industrial diversity in terms of the number of four-digit Japanese SIC manufacturing industries against the population size of UAs in 2000, where all values are in shares again.[60] Comparing UAs in terms of the diversity in IPC subclass and SIC four-digit industry categories is reasonable, since they are comparable in the total number of active categories, which is 608 for the former and 562 for the latter. The solid and dashed lines indicate the fitted OLS lines for the patent class and industrial diversity plots, respectively. While diversity is increasing in the population size of a UA for both patent categories and manufacturing industries, the former is substantially more so: doubling the population size of a UA almost doubles the diversity in the technological category of patents applied in the UA, whereas it only increases the industrial diversity by 55%. Thus, while a larger UA is associated with both larger intensive (i.e., per-capita output) and extensive margins (i.e., diversity) in both R&D and production activities, this tendency is substantially stronger for the former.

These findings are suggestive of a positive association between population concentration and matching externalities promoting collaborator recombinations in large cities.[61] However, the mechanism behind the difference between R&D and industrial location patterns has not been fully explored either theoretically or empirically, and this remains a future research subject.[62]

[58]The location of the patent is identified by the location of the patent applicant. Manufacturing output is obtained from the micro data of the Census of Manufacturers in 2000.

[59]Estimated elasticities of patent output with respect to UA population are similar among alternative output measures: Under IPC subclass and cited count, they are 1.516 and 1.458, respectively.

[60]The industrial diversity of a given UA is defined as the number of four-digit manufacturing industries that have positive employment in the UA.

[61]See, for example, Agrawal et al. (2017); Perlman (2016); Mori and Takeda (2018) for recent empirical studies on geographic agglomeration of R&D activities. In particular, Mori and Takeda (2018) found that the nation-wide development of high-speed railway network had a substantially larger positive impact on the agglomeration of R&D activities than on population agglomeration.

[62]It is also possible to ask if there is any particularly relevant geographic scope of collaborations, e.g., within an establishment, a district, a metropolitan area and an island, and so on. See Gordon (2013) for evidence on the geographic scope of co-authorship in academic research.

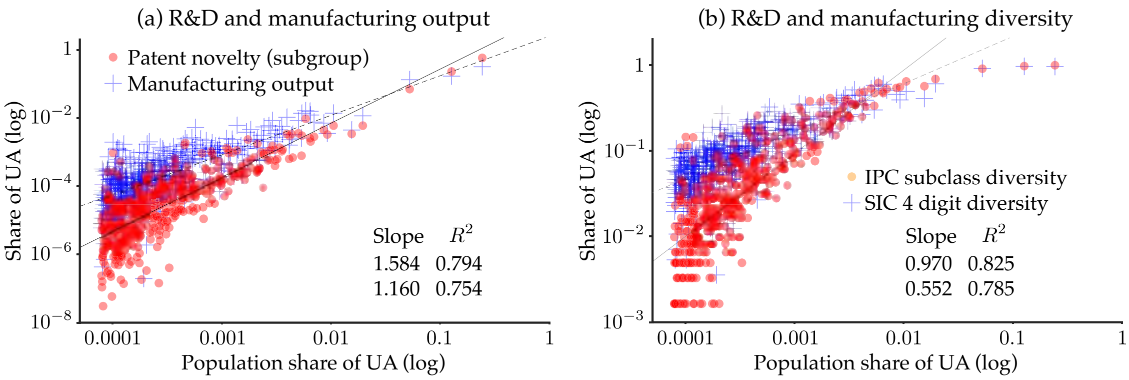


Figure 10.1: Industrial and research outputs and diversities in UAs in 2000

References

Abrams, David S., Ufuk Akcigit, and Jillian Popadak, "Patent value and citations: Creative destruction or strategic disruption?," 2013. Discussion paper No. 19647, National Bureau of Economic Research.

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William R. Kerr, "Innovation, reallocation, and growth," 2017. Discussion Paper No. 12449, Center for Economic Policy Research.

Aghion, Philippe and Peter Howitt, "A model of growth through creative destruction," *Econometrica*, 1992, *60* (2), 323–351.

Agrawal, Ajay, Alberto Galasso, and Alexander Qettl, "Roads and innovation," *Review of Economics and Statistics*, 2017, *99* (3), 417–434.

Akcigit, Ufuk and William R. Kerr, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

_ , Douglas Hanley, and Nicolas Serrano-Velarde, "Back to basic: Basic research spillovers, innovation policy and growth," 2016. Discussion paper No. 11707, Center for Economic Policy Research.

_ , John Grigsby, and Tom Nicholas, "The rise of American ingenuity: Innovation and inventors of the Golden Age," 2017. Discussion paper No. 23047, National Bureau of Economic Research.

_ , Santiago Caicedo, Ernest Miguelez, Stefanie Santcheva, and Valerio Sterzi, "Dancing with the stars: Innovation through interactions," 2018. Discussion paper No. 24466, National Bureau of Economic Research.

Bartolucci, Cristian and Francesco Devicienti, "Better workers move to better firms: A simple test to identify sorting," 2013. Discussion paper No. 7601, Institute of Labor Economics (IZA).

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

Berliant, Marcus and Masahisa Fujita, "Knowledge creation as a square danc on the Hilbert cube," *International Economic Review*, 2008, *49* (4), 1251–1295.

_ and _ , "The Dynamics of Knowledge Diversity and Economic Growth," *Southern Economic Journal*, 2011, *77* (4), 856–884.

_ and _ , "Culture and diversity in knowledge creation," *Regional Science and Urban Economics*, 2012, *42* (4), 648–662.

_ and Tomoya Mori, "Beyond urban form: How Masahisa Fujita shapes us," *International Journal of Economic Theory*, 2017, *13*, 5–28.

Bettencourt, Luís M. A., José Lobo, Dirk Helbing, Cristian Kühnert, and Geoffrey B. West, "Growth, innovation, scaling, and the pace of life in cities," *Proceedings of the National Academy of Sciences of the United States of America*, 2007, *104* (7), 7301–7306.

Boldrin, Michele and David K. Levine, "The case against patents," *Journal of Economic Perspectives*, 2013, *27* (1), 3–22.

Bramoullé, Yann and Bernard Fortin, "Social networks: Econometrics," in Steven N. Durlauf and Lawrence E. Blume, eds., *New Palgrave Dictionary of Economics*, Palgrave Macmillan, 2010.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of peer effects through social networks," *Journal of Econometrics*, 2009, *150* (1), 41–55.

Breschi, Stefano, Francesco Lissoni, and Franco Malerba, "Knowledge-relatedness in firm technological diversification," *Research Policy*, 2003, *32* (1), 69–87.

Calvó-Armengol, Antoni, Elenora Patacchini, and Yves Zenou, "Peer effects and social network in education," *Review of Economic Studies*, 2009, *76* (4), 1239–1267.

Chang, Howard F., "Patent scope, antitrust policy, and cumulative innovation," *RAND Journal of Economics*, 1995, *26* (1), 34–57.

Chu, Angus C., "Effects of blocking patents on R&D: A quantitative DGE analysis," *Journal of Economi Growth*, 2009, *14* (1), 55–78.

Coe, Davit T. and Elhanan Helpman, "International R&D spillovers," *European Economic Review*, 1995, *39* (5), 859–887.

Comola, Margherita and Silvia Prina, "Do interventions change the network: A panel peer-effect model accounting for endogenous network changes," 2014. IZA Discussion paper No.8641.

Cozzi, Guido and Silvia Galli, "Sequential R&D and blocking patents in the dynamics of growth," *Journal of Economic Growth*, 2014, *19* (2), 183–219.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Spatial wage disparities – Workers, firms, and assortative matching," 2016. Unpublished manuscript.

Davis, Donald R. and Jonathan I. Dingel, "The comparative advantage of cities," March 2017. Unpublished manuscript.

_ **and** _ , "A spatial knowledge economy," May 2018. Unpublished manuscript, Columbia University.

Duranton, Gilles and Diego Puga, "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 2001, *91* (5), 1454–1477.

Eeckhout, Jan and Philipp Kircher, "Assortative matching with large firms," *Econometrica*, 2018, *86* (1), 85–132.

Garcia-Vega, Maria, "Does technological diversification promote innovation? An empirical analysis for European firms," *Research Policy*, 2006, *35* (2), 230–246.

Gaubert, Cecile, "Firm sorting and agglomeration," *American Economic Review*, 2018, *forthcoming.*

Giorgi, Giacomo De, Michele Pellizza, and Silvia Redaelli, "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2010, *2* (2), 241–275.

Glass, Amy Jocelyn and Kamal Saggi, "Licensing versus direct investment: Implications for economic growth," *Journal of International Economics*, 2002, *56* (1), 131–153.

Goldsmith-Pinkham, Paul and Guido W. Imbens, "Social networks and identification of peer effects," *Journal of Business & Economic Statistics*, 2013, *31* (3), 253–264.

Gordon, Peter, "Thinking about economic growth: Cities, networs, creativity and supply chains for ideas," *Annals of Regional Science*, 2013, *50* (3), 667–684.

Griliches, Zvi, "Research expenditures, education, and the aggregate agricultural production function," *American Economic Review*, 1964, *54* (6), 961–974.

_ , "Issues in assessing the contribution of research and development to productivity growth," *The Bell Journal of Economics*, 1979, *10* (1), 92–116.

Grossman, Gene M. and Carl Shapiro, "Dynanic R&D competition," *The Economic Journal*, 1978, *97* (386), 372–387.

_ **and Elhanan Helpman**, *Innovation and Growth in the Global Economy*, Cambridge, MA: The MIT Press, 1991.

_ **and** _ , "Quality ladders in the theory of growth," *Review of Economic Studies*, 1991, *58* (1), 43–61.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

Horii, Ryo, "Wants and past knowledge: Growth cycles iwth emerging industries," *Journal of Economic Dynamics and Control*, 2012, *36* (2), 220–238.

Hsieh, Chih-Sheng and Lung-Fei Lee, "A social interactions model with endogenous friendship formation," *Journal of Applied Econometrics*, 2016, *31* (2), 1–21.

Huo, Dong and Kazuyuki Motohashi, "Understanding two types of technological diversity and their effects on the technological value of outcomes from bilateral inter-firm R&D alliances," 2015. Discussion paper No. 15-E-06, The Research Institute of Economy, Trade and Industry (RIETI).

Inoue, Hiroyasu, Kentaro Nakajima, and Yukiko Saito Umeno, "Innovation and collaboration patterns between research establishments," 2015. Discussion paper No. 15-E-049, The Research Institute of Economy, Trade and Industry (RIETI).

_, _, and Yukiko Umeno Saito, "Localization of knowledge-creating establishments," 2016. HIT-REFINED Working Paper Series 47, Institute of Economic Research, Hitotsubashi University.

Inui, Tomohiko, Makiko Nakamuro, Kazuma Edamura, and Junko Ozawa, "Impact of diversity and work-life balance (in Japanese)," 2014. Discussion paper No. 14-J-055, Research Institute of Economy, Trade and Industry (RIETI).

Jackson, Matthew O., Brian W. Rogers, and Yves Zenou, "The economic consequence of social-network structure," *Journal of Economic Literature*, 2017, *55* (1), 49–95.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," *The Quarterly Journal of Economics*, 1993, *108* (3), 577–598.

Jell, Florian, Joachim Henkel, and Martin W. Wallin, "Offensive patent portfolio races," *Long Range Planning*, October 2017, *50* (5), 531–549.

Johnsson, Ida and Hyungsik Roger Moon, "Estimation of peer effects in endogenous social networks: Control function approach," September 2017. Working paper No. 17-25, USC Dornsife Institute for New Economic Thinking.

Jovanovic, Boyan and Rafael Rob, "The diffusion and growth of knowledge," *Review of Economic Studies*, 1989, *56* (4), 569–582.

Kerr, William R. and Scott Duke Kominers, "Agglomerative forces and cluster shapes," *Review of Economics and Statistics*, 2015, *97* (4), 877–899.

Klette, Tor Jakob and Samuel Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

König, Michael D., Xiaodong Liu, and Yves Zenou, "R&D networks: Theory, empirics and policy implications," 2014. Discussion paper No.9872, Center for Economic Policy Research.

Kortum, Samuel, "Research, patenting, and technological change," *Econometrica*, November 1997, *65* (6), 1389–1419.

Laboratory, Inc. Artificial Life, "Patent database for research," January 2018.

Lentz, Rasmus and Dale T. Mortensen, "An empirical model of growth through product innovation," *Econometrica*, November 2008, *76* (6), 1317–1373.

Li, Tong and Li Zhao, "A partial identification subnetwork approach to discrete games in large networks: An application to quantifying peer effects," 2016. Unpublished manuscript, Vanderbilt University.

Manski, Charles F., "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Matutes, Carmen, Pierre Regibeau, and Katharine Rockett, "Optimal patent design and the diffusion of innovations," *RAND Journal of Economics*, 1996, *27* (1), 60–83.

Mendes, Rute, Gerard J. van den Berg, and Maarten Lindeboom, "An empirical assessment of assortative matching in the labor market," *Labour Economics*, 2010, *17* (6), 919–929.

Ministry of Economy, Trade and Industry of Japan, *Census of Manufactures* 1995, 2000, 2005.

__ , *METI Basic Survey of Japanese Business Structure and Activities* 1995-2010.

Ministry of Internal Affairs and Communications of Japan, *Population Census (Tabulation for standard area mesh)* 1995, 2000, 2005.

__ , *Establishment and Enterprise Census* 1996, 2001, 2006.

__ , *Survey of Research and Development* 1997-2010.

__ , *Economic Census for Business Frame* 2009.

__ , *Population Census (Tabulation for standard area mesh)* 2010.

Mori, Tomoya and Alessandro Turrini, "Skills, agglomeration and segmentation," *European Economic Review*, January 2005, *49* (1), 201–225.

__ **and Kohei Takeda**, "Highways, high-speed railways and the growth of Japanese cities," 2018. Unpublished manuscript, Kyoto University.

Murata, Yasusada, Ryosuke Okamoto Ryo Nakajima, and Ryuichi Tamura, "Localized knowledge spillovers and patent citations: A distanced-based approach," *Review of Economics and Statistics*, 2014, *96* (5), 967–985.

Nakajima, Kentaro, Yukiko Umeno Saito, and Iichiro Uesugi, "Localization of interfirm transaction relationships and industry agglomeration," April 2012. Discussion paper 12-E-023, Research Institute of Economy, Trade and Industry.

Nicholas, Tom, "Are patents creative or destructive?," *Antitrust Law Journal*, 2014, *79* (2), 405–421.

Olsson, Ola, "Knowledge as a set in idea space: An epistemological view on growth," *Journal of Economic Growth*, September 2000, *5* (3), 253–275.

__ , "Technological opportunity and growth," *Journal of Economic Growth*, March 2005, *10* (1), 31–53.

Østergaard, Christian R., Bram Timmermans, and Kari Kristinsson, "Does a different view create something new? The effect of employee diversity on innovation," *Research Policy*, 2011, *40* (3), 500–509.

Otazawa, Toshimori, Yuki Ohira, and Jos van Ommeren, "Inter-firm transaction networks and location in a city," August 2018. Discusssion paper 18-E-054, Research Institute of Economy, Trade and Industry.

Panebianco, Fabrizio, Thierry Verdier, and Yves Zenou, "Innovation, pricing and targeting in networks," 2016. Discussion paper No.11398, Center for Economic Policy Research.

Patacchini, Eleonora, Edoardo Rainone, and Yves Zenou, "Heterogeneous peer effects in education," *Journal of Economic Behavior & Organization*, 2017, *134* (C), 190–227.

Perlman, Elizabeth Ruth, "Dense enough to be brilliant: Patents, urbanization, and transportation in nineteenth century America market access, and information flows," January 2016. Unpublished manuscript, Boston University.

Romer, Paul M., "Endogenous technological change," *Journal of Political Economy*, October 1990, *98* (5), 71–102.

Scherer, Frederic M., "Inter-industry technology flows and productivity growth," *Review of Economics and Statistics*, 1982, *64* (4), 627–634.

Schotchmer, Suzanne, "Protecting early inventors: Should second-generation products be patentable?," *RAND Journal of Economics*, 1996, *27* (2), 322–331.

Shell, Karl, "Toward a theory of inventive activitgy and capital accumulation," *American Economic Review*, March 1966, *56* (1/2), 62–68.

_ , *A model of inventive activity and capital accumulation*, Cambridge, MA: The MIT Press,

Tanaka, Hitoshi, "Dynamic analysis of imitation and technology gap," *Journal of Economics*, 2006, *87* (3), 209–240.

_ **, Tatsuro Iwaisako, and Koichi Futagami**, "Dynamic analysis of innovation and international transfer of technology through licensing," *Journal of International Economics*, September 2007, *73* (1), 189–212.

Thompson, Peter and Melanie Fox-Kean, "Patent citations and the geography of knowledge spillovers: A reassessment," *The American Economic Review*, 2005, *95* (1), 450–460.

Tokyo Shoko Research, "Kigyo Zaimu Joho and Sokan File," 2014.

Trajitenberg, Manuel, "A penny for your quotes: Patent citations and the value of innovations," in Adam B. Jaffe and Manuel Trajtenberg, eds., *Patents, Citations & Innovations*, Cambridge, MA: The MIT Press, 2002, chapter 2, pp. 27–49.

Ulku, Hulya, "R&D, innovation, and growth: Evidence from four manufacturing sectors in OECD," *Oxford Economic Papers*, July 2007, *59* (3), 513–535.

Weitzman, Martin L., "Recombinant growth," *Quarterly Journal of Economics*, 1998, *113* (2), 331–360.

Yang, Guifang and Keith E. Maskus, "Intellectual property rights, licensing, and innovation in an endogenous product-cycle model," *Journal of International Economics*, February 2001, *53* (1), 169–187.

Appendix

A Average pairwise productivity of an inventor

To understand pairwise productivity à la Berliant and Fujita (2008), consider two groups of inventors. In group A, two inventors together produce two patents, while in group B, three inventors together produce three patents. For simplicity, let $g_j = 1$ for all patents j. It follows that the proportion of output in each project accruing to one inventor is one-half in group A and one-third in group B. The total output of inventor i is then $\bar{y}_i = 1/2 \times 2 = 1$ in group A and $\bar{y}_i = 1/3 \times 3 = 1$ in group B. However, we assume that knowledge is always created in pairs as in the Berliant-Fujita model. For an inventor in group A, his or her share (one-half) of a given patent is an outcome of the pairwise collaborations with his or her only collaborator; in other words, the proportion of the output of a pairwise collaboration in a given project accruing to him or her is one-half (= $1/2 \div 1$). Since group A produces two patents, the total pairwise output is given by $y_i = 1/2 \times 2 = 1$ for each inventor i. Since an inventor in group B has two collaborators, the proportion of the output of the pairwise collaboration accruing to an inventor for each patent is one-sixth (= $1/3 \div 2$), and the total pairwise output for each inventor i is $1/6 \times 3 = 1/2$. Thus, inventors in group A are more productive in pairwise collaborations than those in group B.

B Locational factors

In this section, the description of UAs and precise definitions for the measures of the local factors discussed in Section 6.3 are given.

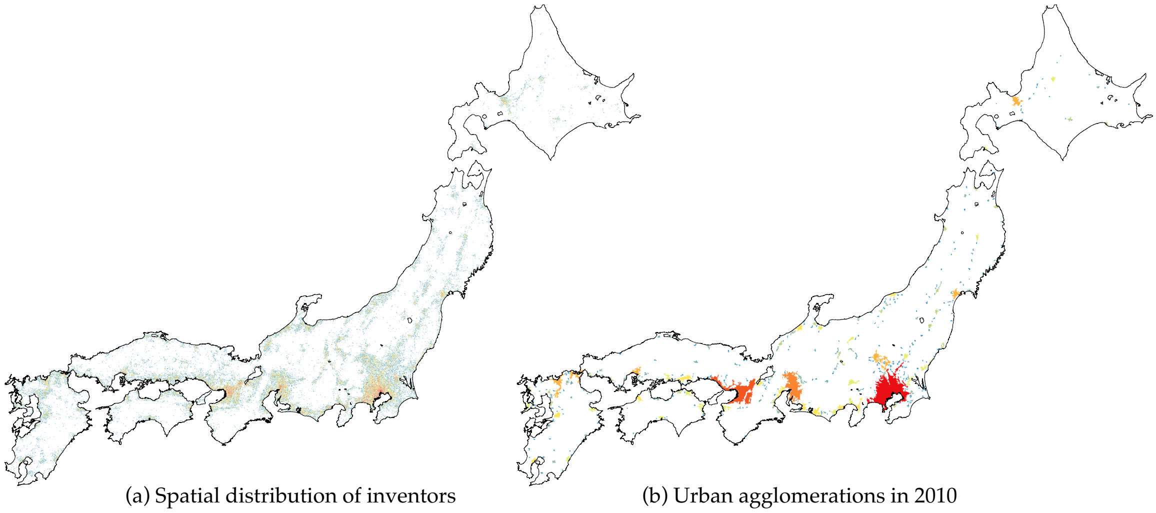
UAs – Panels (a) and (b) in Figure B.1 show the spatial distribution of inventors in I and 453 UAs as of 2010, respectively, where the warmer colors in each panel indicate higher population density. Each inventor is assigned to the closest UA if there is any UA within 10 km of his or her location.

Inventor population – The local population, a_{it}^{INV}, of inventors within a given distance, $\bar{d}$, of the location of inventor i is defined as

$$a_{it}^{\mathrm{INV}} = \left|\left\{j \in I_t \backslash N_{it} \ : \ d(i,j) < \bar{d}\right\}\right|, \tag{B.1}$$

where $d(i,j)$ represents the great-circle distance between inventors i and j (rows 1-4, Table B.1). To evaluate the pure spillover effects, this population excludes the collaborators, N_{it}, of i.[63]

[63]The effects of externalities from the nearby inventors and firms that have been recognized in



(a) Spatial distribution of inventors (b) Urban agglomerations in 2010

Figure B.1: Spatial distribution of researchers and UAs

R&D expenditure – Focusing on manufacturing, we first aggregate firm-level R&D expenditure at the industry level according to the three-digit Japanese SIC in each period t. Denote the industry-level R&D expenditure (in million yen) by v_m for each industry $m \in M_t$, where M_t is the set of three-digit manufacturing industries in period t.[64]

Next, from the micro data of the Establishment and Enterprise Census as well as the Economic Census (MIAC, 1996, 2001, 2006; 2009), we find the set of establishments, E_{mt}, in each industry $m \in M_t$ in period t, and compute the employment share, e_{kt}, of each establishment $k \in E_{mt}$ within industry m.

the literature (e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Murata et al., 2014; Kerr and Kominers, 2015).

[64]Data on R&D expenditure at the firm level are available for firms with at least four employees for every year from 1997 to 2009 from the Survey of Research and Development. Since we do not have data in 1995 and 1996, the total expenditure in 1997–1999 has been inflated by 1.67 times to obtain the value of R&D expenditure in period 0.

Table B.1: Descriptive statistics of the locational factors

Period		(1) 1	(2) 2
(1) inventor population	1km	5,750 (7,225)	5,629 (7,282)
(2)	5km	31,026 (42,143)	30,158 (42,269)
(3)	10km	70,720 (79,277)	66,011 (77,330)
(4)	20km	140,204 (129,401)	127,470 (120,751)
(5) R&D investment	1km	10,454 (78,020)	18,480 (180,284)
(6)	5km	150,581 (338,668)	278,911 (703,381)
(7)	10km	300,256 (466,130)	520,066 (920,505)
(8)	20km	550,420 (584,891)	899,652 (1,098,091)
(9) Manufacturing employment	1km	2,240 (1,505)	6,676 (7,106)
(10)	5km	52,974 (32,395)	76,491 (74,655)
(11)	10km	182,597 (106,414)	212,371 (166,473)
(12)	20km	551,875 (318,789)	509,703 (322,326)
(13) Manufacturing output (in thousand)	1km	21,801,942 (58,182,730)	20,774,589 (83,883,736)
(14)	5km	158,183,183 (129,167,825)	104,957,604 (129,388,708)
(15)	10km	445,908,195 (255,976,915)	317,846,559 (226,259,080)
(16)	20km	1,213,122,353 (626,842,420)	956,808,207 (532,719,932)
(17) Residential population	5km	595,461 (386,442)	615,722 (399,930)
(18)	10km	2,100,541 (1,388,078)	2,156,271 (1,432,171)
(19)	20km	6,386,959 (4,252,098)	6,573,357 (4,416,168)

Numbers in parentheses are standard deviations.

Assuming that the R&D expenditure of each establishment in each industry is proportional to the employment size of the establishment, the value of R&D expenditure of each establishment in period t is approximated by $v_{mt}e_{mt}$. Assuming that the R&D expenditure in the previous period $t-1$ affects the productivity of inventors in the current period t, the R&D around inventor i in period t is given as follows (rows 5-8, Table B.1):[65]

$$a_{it}^{\text{R\&D}} = \sum_{m \in M_t} \sum_{k \in \{j \in E_m : d(i,j) < \bar{d}\}} v_{m,t-1} e_{k,t-1}. \tag{B.2}$$

[65]The R&D expenditure values are obtained from the Survey of Research and Development (1997-2010b) by MIAC and from METI Basic Survey of Japanese Business Structure and Activities (1995-2010) by METI.

$a_{it}^{\text{R\&D}}$ naturally influences patent development (e.g., Griliches, 1979; Coe and Helpman, 1995; Ulku, 2007).

Manufacturing concentration – Assuming that the employment size and output of an establishment correlate with demand for new knowledge, we proxy the local market size for an invented technology around inventor i by the local manufacturing employment and output around i:[66]

$$a_{it}^{\text{MNF}_j} = \sum_{k\in\{j\in E_t : d(i,j)<\bar{d}\}} e_{kt} \tag{B.3}$$

where $E_t = \cup_{m\in M_t} E_{mt}$, and e_{kt} represents the total output value (employment) of establishment k for $j = o$ ($j = e$) (rows 9-16, Table B.1).[67]

Residential population – The local residential population is defined as

$$a_{it}^{\text{POP}} = \sum_{k\in\{j\in R : d(i,j)<\bar{d}\}} r_{kt} \tag{B.4}$$

where R represents the set of 1km-by-1km cells covering the relevant location space in Japan; the centroid of each cell is considered to be the representative location of the cell in measuring the distance from the cell; r_{kt} is the residential population in cell $k \in R$ at the beginning of period t (rows 17-19, Table B.1).[68]

C Similarity and difference with linear-in-means models

In this section, we discuss the similarity and difference of instruments between the linear-in-means models of social interactions as in Bramoullé et al. (2009) and our model.

The most relevant similarity is the reflection problem intrinsic to the agent network in both cases, while the most fundamental difference is whether the relevance of the IVs is intrinsic or extrinsic to the network of agents in the model.

In the case of the peer effects in the linear-in-means models, the relevance accrues from the simultaneous equation structure of the model, and thus it is intrinsic to

[66] Another interpretation of a_{it}^{MNF} is the spillover from the manufacturing concentration around i in period t.

[67] The manufacturing employment values are obtained from the Establishment and Enterprise Census for (1996, 2001, 2006) and Economic Census for Business Frame (2009) by MIAC; the manufacturing output values are obtained from the micro data of the Census of Manufacturers (1995, 2000, 2005) and Economic Census for Business Frame (2009) by MIAC.

[68] The residential population in the 1 km-by-1 km cells is available from the Population Census (1995, 2000, 2005) by MIAC.

the network. As a consequence, adding degrees of separation in the network is double-edged: the IVs constructed from more distant indirect collaborators can gain exogeneity only at the cost of loosing the relevance. For this reason, the IVs in Bramoullé et al. (2009) are constructed from the exogenous variables of relatively close indirect collaborators in order to retain sufficiently strong relevance. A great advantage in their model is that their IVs formally satisfy the exclusion restriction, provided that the network is exogenous.

In our case, the relevance of the IVs is extrinsic to the inventor network, since it comes from the similarity in inventor productivity as a result of assortative matching between firms and workers that happened prior to the formation of the inventor network. As a consequence, the relevance is maintained even when the information of the distant indirect collaborators is solely used, as long as the assortative matching affects the indirect collaborators and the targeted inventors simultaneously. That is, the increasing the separation in the network is not double edged. While the endogeneity of the IVs is only virtually (but never completely) eliminated by using sufficiently distant indirect collaborators to construct IVs in our case unlike the case of the linear-in-means models, we instead can allow for the endogenous network formation.

D First-stage regressions

This section presents the results of the first-stage regressions for the 2SLS IV regressions corresponding to columns 2-5 and 7-10 in Table 8.1 and those in Table 8.3 in Tables D.1 and D.2, respectively.

Table D.1: Regression results (Dependent variable: $\ln k_{it}^{D}$)

	Citations				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) k_{it}^{D,IV_3}	0.436***	0.453***			0.340***	0.402***		
	(0.0266)	(0.0169)			(0.0146)	(0.0132)		
(2) k_{it}^{D,IV_4}	0.0235*		0.349***		0.0880***		0.347***	
	(0.0128)		(0.0136)		(0.0149)		(0.0149)	
(3) k_{it}^{D,IV_5}	0.00544			0.249***	0.0411*			0.266***
	(0.0409)			(0.0271)	(0.0225)			(0.0250)
(4) $\ln k_{it}$	0.124***	0.124***	0.131***	0.134***	0.156***	0.156***	0.166***	0.176***
	(0.0212)	(0.0210)	(0.0266)	(0.0304)	(0.0126)	(0.0128)	(0.0130)	(0.0145)
(5) $(\ln k_{it})^2$	-0.0491***	-0.0491***	-0.0524***	-0.0545***	-0.0900***	-0.0892***	-0.0934***	-0.0949***
	(0.00741)	(0.00738)	(0.0107)	(0.0126)	(0.00766)	(0.00777)	(0.00967)	(0.0114)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.359***	0.360***	0.390***	0.402***	0.504***	0.515***	0.538***	0.561***
	(0.0787)	(0.0772)	(0.0845)	(0.0859)	(0.0969)	(0.0988)	(0.108)	(0.115)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00240	0.00259	0.00553	0.00979	0.0140	0.0158	0.0178	0.0252
	(0.00909)	(0.00921)	(0.0104)	(0.0113)	(0.0137)	(0.0142)	(0.0166)	(0.0181)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0668***	-0.0663***	-0.0745***	-0.0759***	-0.0954***	-0.0943***	-0.111***	-0.118***
	(0.0196)	(0.0189)	(0.0244)	(0.0266)	(0.0221)	(0.0197)	(0.0238)	(0.0239)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0227	0.0227	0.0242	0.0242	0.0160	0.0151	0.0149	0.00968
	(0.0207)	(0.0204)	(0.0247)	(0.0246)	(0.0294)	(0.0273)	(0.0320)	(0.0317)
(10) $\ln a_{it}^{\mathrm{POP}}$	1.139	1.143	1.391	1.510	3.041***	3.084***	3.466***	3.699***
	(0.935)	(0.918)	(1.112)	(1.148)	(1.042)	(0.985)	(1.163)	(1.240)
(11) τ_1	0.285***	0.288***	0.333***	0.369***	0.474***	0.504***	0.548***	0.611***
	(0.0211)	(0.0190)	(0.0233)	(0.0288)	(0.0346)	(0.0345)	(0.0382)	(0.0443)
(12) R^2	0.205	0.205	0.183	0.171	0.203	0.201	0.188	0.179
(13) F	443.2	718.4	652.5	84.21	398.6	925.4	541.2	113.2
(14) p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(15) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

Table D.2: Regression results (Dependent variable: $\ln \Delta n_{it}$)

	Citations				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) $\ln \Delta n_{it}^{\mathrm{IV}_3}$	0.244***	0.278***			0.244***	0.278***		
	(0.0212)	(0.0138)			(0.0212)	(0.0138)		
(2) $\ln \Delta n_{it}^{\mathrm{IV}_4}$	0.00997		0.231***		0.00997		0.231***	
	(0.0321)		(0.0304)		(0.0321)		(0.0304)	
(3) $\ln \Delta n_{it}^{\mathrm{IV}_5}$	0.106***			0.208***	0.106***			0.208***
	(0.0339)			(0.0392)	(0.0339)			(0.0392)
(4) $\ln k_{it}$	0.114***	0.116***	0.117***	0.117***	0.114***	0.116***	0.117***	0.117***
	(0.0268)	(0.0266)	(0.0261)	(0.0255)	(0.0268)	(0.0266)	(0.0261)	(0.0255)
(5) $(\ln k_{it})^2$	-0.202***	-0.202***	-0.203***	-0.204***	-0.202***	-0.202***	-0.203***	-0.204***
	(0.00986)	(0.00962)	(0.00971)	(0.00984)	(0.00986)	(0.00962)	(0.00971)	(0.00984)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.232***	0.237***	0.261***	0.273***	0.232***	0.237***	0.261***	0.273***
	(0.0529)	(0.0542)	(0.0600)	(0.0651)	(0.0529)	(0.0542)	(0.0600)	(0.0651)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00704	0.00845	0.00680	0.00712	0.00704	0.00845	0.00680	0.00712
	(0.00751)	(0.00733)	(0.00840)	(0.00874)	(0.00751)	(0.00733)	(0.00840)	(0.00874)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0329	-0.0326*	-0.0412*	-0.0427*	-0.0329	-0.0326*	-0.0412*	-0.0427*
	(0.0200)	(0.0194)	(0.0220)	(0.0240)	(0.0200)	(0.0194)	(0.0220)	(0.0240)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.00479	0.00514	0.00793	0.00868	0.00479	0.00514	0.00793	0.00868
	(0.0142)	(0.0142)	(0.0155)	(0.0170)	(0.0142)	(0.0142)	(0.0155)	(0.0170)
(10) $\ln a_{it}^{\mathrm{POP}}$	1.053**	1.106**	1.210**	1.338**	1.053**	1.106**	1.210**	1.338**
	(0.506)	(0.539)	(0.585)	(0.585)	(0.506)	(0.539)	(0.585)	(0.585)
(11) τ_1	-0.167***	-0.146***	-0.139***	-0.131***	-0.167***	-0.146***	-0.139***	-0.131***
	(0.0206)	(0.0187)	(0.0252)	(0.0263)	(0.0206)	(0.0187)	(0.0252)	(0.0263)
(12) R^2	0.197	0.196	0.190	0.189	0.197	0.196	0.190	0.189
(13) F	142.2	406.1	57.40	28.24	142.2	406.1	57.40	28.24
(14) p-value	0.000	0.000	0.000	3.44e-07	0.000	0.000	0.000	3.44e-07
(15) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

E Results for model (5.9) with $m = p$

Tables E.1 shows the second stage regression results for model (5.9) with $m = p$.

Table E.1: Regression results (Dependent variable: $\ln y^p_{it}$)

	Quality					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k^D_{it}$	0.135***	0.260***	0.259***	0.272***	0.270***	0.0451***	0.114***	0.110***	0.132***	0.130***
	(0.0103)	(0.0283)	(0.0290)	(0.0271)	(0.0373)	(0.00521)	(0.0214)	(0.0246)	(0.0170)	(0.0416)
(2) $\ln k_{it}$	0.0999***	0.0825***	0.0826***	0.0808***	0.0811***	0.111***	0.0982***	0.0989***	0.0949***	0.0952***
	(0.0110)	(0.00729)	(0.00725)	(0.00844)	(0.0125)	(0.0144)	(0.0112)	(0.0106)	(0.0125)	(0.0194)
(3) $(\ln k_{it})^2$	-0.0835***	-0.0765***	-0.0765***	-0.0758***	-0.0759***	-0.0868***	-0.0803***	-0.0807***	-0.0787***	-0.0788***
	(0.00921)	(0.00761)	(0.00759)	(0.00796)	(0.00884)	(0.0104)	(0.00821)	(0.00800)	(0.00869)	(0.0115)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.207***	0.153**	0.153**	0.148**	0.149***	0.238***	0.196***	0.199***	0.185***	0.186***
	(0.0678)	(0.0752)	(0.0755)	(0.0700)	(0.0574)	(0.0651)	(0.0659)	(0.0683)	(0.0612)	(0.0389)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0298***	0.0281***	0.0281***	0.0279***	0.0280***	0.0302***	0.0280***	0.0282***	0.0275***	0.0276***
	(0.0105)	(0.00925)	(0.00927)	(0.00892)	(0.00867)	(0.0112)	(0.0101)	(0.0102)	(0.00967)	(0.00885)
(6) $\ln a^{\mathrm{MNF}e}_{it}$	-0.0122*	-0.00309	-0.00313	-0.00222	-0.00236	-0.0166**	-0.00823	-0.00872	-0.00604	-0.00626
	(0.00702)	(0.00616)	(0.00611)	(0.00725)	(0.00869)	(0.00772)	(0.00745)	(0.00756)	(0.00718)	(0.00950)
(7) $\ln a^{\mathrm{MNF}o}_{it}$	-0.000286	-0.00341	-0.00340	-0.00371	-0.00366	0.00274	0.00217	0.00220	0.00202	0.00203
	(0.00354)	(0.00438)	(0.00438)	(0.00435)	(0.00426)	(0.00420)	(0.00344)	(0.00345)	(0.00339)	(0.00343)
(8) $\ln a^{\mathrm{POP}}_{it}$	0.133	-0.0794	-0.0785	-0.0997	-0.0965	0.183	-0.0932	-0.0768	-0.165	-0.158
	(0.461)	(0.420)	(0.420)	(0.409)	(0.387)	(0.535)	(0.509)	(0.517)	(0.495)	(0.458)
(9) τ_1	0.126***	0.0708***	0.0710***	0.0655***	0.0663***	0.153***	0.103***	0.106***	0.0898***	0.0911***
	(0.0183)	(0.0215)	(0.0217)	(0.0199)	(0.0223)	(0.0179)	(0.0236)	(0.0254)	(0.0206)	(0.0342)
(10) R^2	0.102	0.087	0.087	0.084	0.084	0.089	0.078	0.079	0.072	0.072
(11) Hansen J p-val.		0.878					0.177			
(12) 1st stage F		727.1	2178	1080	509.6		557.6	1590	918.7	471.4
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

F Results under alternative IVs

Table F.1: Regression results for (5.1) under alternative IVs

	Quality					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k^D_{it}$	0.166***	0.225***	0.240***	0.0980	-0.0272	0.167***	0.284***	0.288***	0.268***	0.291***
	(0.0106)	(0.0455)	(0.0404)	(0.111)	(0.191)	(0.00529)	(0.0412)	(0.0418)	(0.0444)	(0.0702)
(2) $\ln k_{it}$	0.115***	0.107***	0.105***	0.125***	0.142***	0.155***	0.134***	0.133***	0.137***	0.133***
	(0.0165)	(0.0198)	(0.0183)	(0.0339)	(0.0497)	(0.0169)	(0.0237)	(0.0236)	(0.0242)	(0.0270)
(3) $(\ln k_{it})^2$	-0.0887***	-0.0855***	-0.0847***	-0.0924***	-0.0992***	-0.194***	-0.183***	-0.183***	-0.185***	-0.183***
	(0.0105)	(0.0123)	(0.0118)	(0.0174)	(0.0231)	(0.00999)	(0.00675)	(0.00641)	(0.00742)	(0.0119)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.187***	0.161***	0.154**	0.217***	0.272***	0.334***	0.262***	0.259***	0.272***	0.258***
	(0.0603)	(0.0613)	(0.0618)	(0.0684)	(0.0919)	(0.0927)	(0.0940)	(0.0966)	(0.0912)	(0.0582)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0279***	0.0274***	0.0273***	0.0285***	0.0295***	0.0434***	0.0403***	0.0402***	0.0407***	0.0401***
	(0.00781)	(0.00734)	(0.00723)	(0.00838)	(0.00945)	(0.0151)	(0.0132)	(0.0132)	(0.0133)	(0.0121)
(6) $\ln a^{\mathrm{MNF}e}_{it}$	0.0117*	0.0156***	0.0166***	0.00724	-0.00102	-0.0130	0.00138	0.00189	-0.000597	0.00222
	(0.00695)	(0.00576)	(0.00559)	(0.00774)	(0.0100)	(0.0115)	(0.00923)	(0.00884)	(0.00978)	(0.0164)
(7) $\ln a^{\mathrm{MNF}o}_{it}$	0.00630	0.00492	0.00457	0.00788	0.0108*	-0.00250	-0.00317	-0.00319	-0.00308	-0.00321
	(0.00670)	(0.00716)	(0.00755)	(0.00481)	(0.00616)	(0.00642)	(0.00801)	(0.00808)	(0.00778)	(0.00804)
(8) $\ln a^{\mathrm{POP}}_{it}$	-0.644	-0.713	-0.731	-0.565	-0.419	0.818	0.405	0.390	0.462	0.381
	(0.533)	(0.514)	(0.508)	(0.601)	(0.731)	(0.503)	(0.495)	(0.504)	(0.488)	(0.411)
(9) τ_1	0.225***	0.201***	0.195***	0.253***	0.306***	0.303***	0.222***	0.219***	0.233***	0.217***
	(0.0168)	(0.0257)	(0.0231)	(0.0562)	(0.0934)	(0.0346)	(0.0545)	(0.0550)	(0.0557)	(0.0610)
(10) R^2	0.153					0.186				
(11) Hansen J p-val.		0.494					0.373			
(12) 1st stage F		322.6	944.7	522.2	209		372.1	1041	694.2	387.1
(13) #Obs.	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table F.2: Regression results for (5.10) under alternative IVs

	Quality					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln \Delta n_{it}$	0.0993***	0.900***	0.901***	0.956***	0.986***	0.236***	1.216***	1.217***	1.280***	1.316***
	(0.00687)	(0.0293)	(0.0299)	(0.0352)	(0.0301)	(0.00901)	(0.0482)	(0.0491)	(0.0621)	(0.0611)
(2) $\ln k_{it}$	0.133***	0.0365	0.0363	0.0296	0.0260	0.152***	0.0341	0.0339	0.0263	0.0221
	(0.0423)	(0.0562)	(0.0562)	(0.0567)	(0.0580)	(0.0328)	(0.0483)	(0.0482)	(0.0480)	(0.0492)
(3) $(\ln k_{it})^2$	-0.0390***	0.123***	0.123***	0.135***	0.141***	-0.0458***	0.153***	0.153***	0.166***	0.173***
	(0.0148)	(0.0162)	(0.0160)	(0.0153)	(0.0165)	(0.0142)	(0.0135)	(0.0133)	(0.0114)	(0.0127)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.376***	0.146***	0.146***	0.130***	0.121***	0.501***	0.220***	0.219***	0.201**	0.191**
	(0.0928)	(0.0354)	(0.0352)	(0.0324)	(0.0319)	(0.114)	(0.0777)	(0.0779)	(0.0794)	(0.0778)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0100	0.00329	0.00328	0.00281	0.00256	0.0290*	0.0208**	0.0208**	0.0202**	0.0199**
	(0.0109)	(0.00561)	(0.00560)	(0.00541)	(0.00528)	(0.0165)	(0.0102)	(0.0102)	(0.01000)	(0.00984)
(6) $\ln a_{it}^{\mathrm{MNFe}}$	-0.0680**	-0.0322**	-0.0321**	-0.0296**	-0.0283*	-0.109***	-0.0653***	-0.0653***	-0.0624***	-0.0608***
	(0.0269)	(0.0151)	(0.0151)	(0.0149)	(0.0146)	(0.0219)	(0.0150)	(0.0150)	(0.0154)	(0.0159)
(7) $\ln a_{it}^{\mathrm{MNFo}}$	0.0210	0.0118	0.0118	0.0112	0.0108	0.000656	-0.0106	-0.0106	-0.0113	-0.0118
	(0.0231)	(0.0115)	(0.0115)	(0.0109)	(0.0107)	(0.0266)	(0.0106)	(0.0106)	(0.00974)	(0.00940)
(8) $\ln a_{it}^{\mathrm{POP}}$	1.018	-0.161	-0.163	-0.245	-0.289	3.127***	1.683*	1.681*	1.588*	1.536
	(1.124)	(1.130)	(1.129)	(1.136)	(1.148)	(1.194)	(0.960)	(0.959)	(0.945)	(0.950)
(9) τ_1	0.393***	0.456***	0.456***	0.460***	0.463***	0.668***	0.745***	0.746***	0.751***	0.753***
	(0.0325)	(0.0443)	(0.0443)	(0.0455)	(0.0457)	(0.0312)	(0.0269)	(0.0269)	(0.0279)	(0.0280)
(10) R^2	0.159					0.175				
(11) Hansen J p-val.		0.194					0.185			
(12) 1st stage F		2800	8401	6651	5010		2800	8401	6651	5010
(13) #Obs.	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

G Results under alternative productivity measures

This section presents the regression results for (5.1) and (5.10) under the four alternative measures of inventor productivity, where the output, g_j, of patent j in (2.1) is given by (*i*) the cited count within five years from publication, (*ii*) technological novelty based on the IPC subclass, (*iii*) count of patent claims; or (*iv*) count of patents, i.e., $g_j = 1$ for all j. Table G.1 shows the descriptive statistics for productivities and differentiated knowledge of collaborators under these measures.

Table G.1: Descriptive statistics of knowledge and productivity variables

Unit of productivity		**Citations (5 years)**		**Novelty (IPC subclass)**		**Claim counts**		**Patent counts**	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period		1	2	1	2	1	2	1	2
(1) Output of a patent	g_{jt}	1.789	1.595	0.000	0.000	7.231	8.906	1.000	1.000
		(3.676)	(4.186)	(0.002)	(0.003)	(9.555)	(81.53)	(0.000)	(0.000)
(2) Productivity of an inventor	y_{it}	9.369	5.597	0.001	0.000	36.67	40.89	4.824	3.099
		(22.02)	(163.73)	(0.004)	(0.003)	(109.14)	(4173.48)	(7.749)	(5.936)
(3) Pairwise productivity of an inventor	y_{it}	1.622	1.838	0.000	0.000	6.682	25.48	0.894	0.677
		(3.730)	(175.40)	(0.001)	(0.001)	(88.27)	(4478.60)	(1.911)	(1.739)
(4) Avg. diff. knowledge of collaborators	k_{it}^D	1.666	1.186	0.008	0.005	6.579	5.647	0.874	0.748
		(8.940)	(5.391)	(0.043)	(0.034)	(49.51)	(25.86)	(4.699)	(3.057)

Numbers in parentheses are standard deviations.

Tables G.2 and G.3 present the results from the second-stage regressions for models (5.1) and (5.10), respectively. Under each alternative measure, the tables show the OLS and the IV results, where the IVs for $\ln k_{it}^D$ and $\ln \Delta n_{it}$ are constructed by using all indirect collaborators for $\ell = 3,4$ and 5, since the result is similar if only one of them is used (just like in our baseline results).

Table G.2: Regression results for (5.1) under alternative productivity measures

	Citations (5 years)		Novelty (IPC subclass)		Claim count		Patent count	
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5	(7)OLS	(8) IV3-5
(1) $\ln k^D_{it}$	0.163***	0.282***	0.179***	0.323***	0.198***	0.339***	0.163***	0.326***
	(0.0109)	(0.0276)	(0.00882)	(0.0347)	(0.0133)	(0.0383)	(0.0112)	(0.0304)
(2) $\ln k_{it}$	0.116***	0.0991***	0.0586***	0.0394**	0.140***	0.115***	0.0981***	0.0775***
	(0.0155)	(0.0117)	(0.0199)	(0.0186)	(0.0135)	(0.0152)	(0.0114)	(0.00723)
(3) $(\ln k_{it})^2$	-0.0887***	-0.0820***	-0.159***	-0.146***	-0.0953***	-0.0858***	-0.0826***	-0.0742***
	(0.01000)	(0.00892)	(0.0184)	(0.0121)	(0.00337)	(0.00338)	(0.00911)	(0.00715)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.167***	0.118*	0.232***	0.153**	0.212***	0.142**	0.187***	0.109
	(0.0539)	(0.0606)	(0.0644)	(0.0707)	(0.0559)	(0.0579)	(0.0658)	(0.0734)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0269***	0.0254***	0.0415***	0.0379***	0.0276***	0.0251***	0.0290***	0.0265***
	(0.00744)	(0.00650)	(0.0127)	(0.0106)	(0.0101)	(0.00865)	(0.00971)	(0.00767)
(6) $\ln a^{\mathrm{MNF}_e}_{it}$	0.0188***	0.0269***	0.00689	0.0214**	0.0148**	0.0274***	-0.00502	0.0120**
	(0.00566)	(0.00465)	(0.0107)	(0.00993)	(0.00658)	(0.00542)	(0.00639)	(0.00504)
(7) $\ln a^{\mathrm{MNF}_o}_{it}$	0.00857	0.00546	-0.00516	-0.00643	0.0127**	0.00860	0.000798	-0.00151
	(0.00616)	(0.00835)	(0.00456)	(0.00637)	(0.00508)	(0.00604)	(0.00344)	(0.00586)
(8) $\ln a^{\mathrm{POP}}_{it}$	-0.435	-0.626	0.394	0.00176	0.594	0.154	0.0162	-0.331
	(0.527)	(0.494)	(0.533)	(0.466)	(0.478)	(0.484)	(0.450)	(0.409)
(9) τ_1	0.272***	0.214***	0.433***	0.327***	0.122***	0.0788***	0.133***	0.0801***
	(0.0175)	(0.0156)	(0.0247)	(0.0352)	(0.0236)	(0.0230)	(0.0160)	(0.0155)
(10) R^2	0.165		0.236		0.105		0.107	
(11) Hansen J p-val.		0.845		0.184		0.399		0.629
(12) 1st stage F		712.6		500.8		889.2		774.7
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table G.3: Regression results for (5.10) under alternative productivity measures

	Citations (5 years)		Novelty (IPC subclass)		Claim count		Patent count	
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5	(7)OLS	(8) IV3-5
(1) $\ln \Delta n_{it}$	0.107***	1.383***	0.174***	1.427***	0.110***	1.525***	0.0851***	1.318***
	(0.00587)	(0.0683)	(0.00853)	(0.0920)	(0.00630)	(0.0540)	(0.00595)	(0.0512)
(2) $\ln k_{it}$	0.137***	-0.0166	0.122***	0.000294	0.166***	-0.00414	0.116***	-0.0324
	(0.0452)	(0.0688)	(0.0467)	(0.0914)	(0.0234)	(0.0526)	(0.0307)	(0.0563)
(3) $(\ln k_{it})^2$	-0.0366**	0.225***	-0.0551	0.242***	-0.0456***	0.244***	-0.0343***	0.218***
	(0.0160)	(0.0191)	(0.0340)	(0.0389)	(0.0103)	(0.0200)	(0.0131)	(0.0200)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.365***	-0.00994	0.479***	0.144**	0.445***	0.0288	0.450***	0.0876
	(0.0965)	(0.0387)	(0.0967)	(0.0611)	(0.131)	(0.0561)	(0.0913)	(0.0534)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0129	0.000127	0.0239	0.00825	0.0186	0.00446	0.0156	0.00330
	(0.0106)	(0.00535)	(0.0151)	(0.00960)	(0.0129)	(0.00568)	(0.0140)	(0.00545)
(6) $\ln a^{\mathrm{MNF}_e}_{it}$	-0.0666***	-0.00955	-0.0990***	-0.0673***	-0.0916***	-0.0283	-0.106***	-0.0505***
	(0.0212)	(0.0150)	(0.0220)	(0.0176)	(0.0233)	(0.0203)	(0.0219)	(0.0139)
(7) $\ln a^{\mathrm{MNF}_o}_{it}$	0.0221	0.00871	0.00282	-0.0220*	0.0225	0.00768	0.00888	-0.00401
	(0.0204)	(0.0112)	(0.0265)	(0.0130)	(0.0257)	(0.0132)	(0.0253)	(0.00891)
(8) $\ln a^{\mathrm{POP}}_{it}$	1.238	-0.698	2.201**	-0.0572	2.924***	0.778	1.808*	-0.0609
	(1.058)	(1.278)	(0.879)	(1.056)	(0.886)	(1.167)	(0.951)	(1.000)
(9) τ_1	0.459***	0.560***	0.674***	0.522***	0.287***	0.398***	0.301***	0.398***
	(0.0289)	(0.0539)	(0.0265)	(0.0398)	(0.0262)	(0.0322)	(0.0215)	(0.0319)
(10) R^2	0.177		0.217		0.089		0.111	
(11) Hansen J p-val.		0.254		0.297		0.245		0.251
(12) 1st stage F		237.7		251.2		237.7		237.7
(13) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

H Results under the alternative definition of collaborators' knowledge

Table H.1 shows the regression results for (5.1) under quality- and novelty-adjusted productivity measures in columns 1-2 and 3-4, respectively, and those for (5.10) in columns 5-6 when k_{it}^{D} is defined in terms of IPC subgroups as given by (9.1). For each specification, we compare the OLS and IV results, where the latter are shown only for the case in which IVs are constructed by using all the third-, forth- and fifth-indirect collaborators, since the result is similar, even if only either of the third-, forth- or fifth-indirect collaborators were used.

In all the specifications, the first-stage F values are reasonably large, so that the relevance appears to be strong as in the baseline case. In terms of the Hansen (1982)'s J-test, there is no evidence against the exogeneity of the instruments.

Table H.1: Regression results with knowledge in terms of IPC subgroups

	(5.1) Dependent variable : $\ln y_{it}$				(5.10) Dependent variable : $\ln \Delta n_{it}$	
	Quality		Novelty			
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5
(1) $\ln k_{it}^{D}$	0.169***	0.604***	0.287***	1.331***		
	(0.0431)	(0.118)	(0.0597)	(0.244)		
(2) $\ln \Delta n_{it}$					-0.0210***	0.368***
					(0.00386)	(0.0160)
(3) $\ln k_{it}$	0.128***	0.111***	0.167***	0.126***	0.0432***	-0.00226
	(0.0201)	(0.0166)	(0.0163)	(0.0246)	(0.00642)	(0.0157)
(4) $(\ln k_{it})^2$	-0.0983***	-0.0936***	-0.208***	-0.197***	-0.0179***	0.0611***
	(0.0116)	(0.0107)	(0.0106)	(0.00802)	(0.00184)	(0.00651)
(5) $\ln a_{it}^{\mathrm{INV}}$	0.230***	0.189***	0.386***	0.287***	0.105***	-0.00729
	(0.0596)	(0.0610)	(0.103)	(0.109)	(0.0277)	(0.0213)
(6) $\ln a_{it}^{\mathrm{R\&D}}$	0.0281***	0.0270***	0.0448***	0.0422***	0.00259	-0.00116
	(0.00879)	(0.00771)	(0.0170)	(0.0143)	(0.00380)	(0.00188)
(7) $\ln a_{it}^{\mathrm{MNF}_e}$	0.00313	0.00820	-0.0293**	-0.0171	-0.0148***	0.00371
	(0.00823)	(0.00705)	(0.0114)	(0.0110)	(0.00461)	(0.00650)
(8) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0113***	0.0112**	-0.00369	-0.00392	-0.00126	-0.00551
	(0.00431)	(0.00555)	(0.00661)	(0.00810)	(0.00579)	(0.00378)
(9) $\ln a_{it}^{\mathrm{POP}}$	-0.375	-0.727	1.238**	0.392	0.903***	0.299
	(0.627)	(0.635)	(0.535)	(0.587)	(0.200)	(0.247)
(10) τ_1	0.279***	0.256***	0.399***	0.344***	0.0453***	0.0754***
	(0.0195)	(0.0176)	(0.0339)	(0.0387)	(0.00639)	(0.00867)
(11) R^2	0.132		0.169		0.018	
(12) Hansen J p-val.		0.931		0.109		0.235
(13) 1st stage F		328.8		328.8		235.2
(14) #Obs.	113,454	113,454	113,454	113,454	92,098	92,098

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

I Results for the alternative radiuses for locational factors

This section presents the results from the second-stage regressions for (5.1) in Section 8.1 and (5.10) in Section 8.3 under the alternative radius values for the local factors defined in Section 6.3 in Tables I.1 and I.2 (I.3 and I.4), respectively for quality-adjusted (novelty-adjusted) productivity.

One can see that the choice of radius values for the local factors does not alter the qualitative results obtained in the baseline setup shown in Tables 8.1 and 8.3 in Section 8 regarding the effect of collaborators' differentiated knowledge and that of the knowledge stock of an inventor on his or her productivity as well as the role of the collaborator recombination in the size of collaborators' differentiated knowledge. The values of the estimated coefficients for the endogenous variables, $\ln k_{it}^{D}$ and $\ln \Delta n_{it}$, as well as those for the knowledge stock, $\ln k_{it}$ and $(\ln k_{it})^2$, appear to be stable in all cases.

Table I.1: Regression results (Dependent variable: $\ln y_{it}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k_{it}^{D}$	0.293***	0.296***	0.294***	0.286***	0.288***	0.288***	0.286***
	(0.0225)	(0.0193)	(0.0214)	(0.0254)	(0.0239)	(0.0242)	(0.0254)
(2) $\ln k_{it}$	0.0923***	0.0893***	0.0889***	0.0931***	0.0917***	0.0911***	0.0931***
	(0.0124)	(0.0135)	(0.0141)	(0.0119)	(0.0121)	(0.0121)	(0.0119)
(3) $(\ln k_{it})^2$	-0.0814***	-0.0803***	-0.0802***	-0.0820***	-0.0815***	-0.0816***	-0.0820***
	(0.00886)	(0.00903)	(0.00911)	(0.00868)	(0.00879)	(0.00866)	(0.00868)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km				0.117*	0.126**	0.125*	0.117*
				(0.0633)	(0.0634)	(0.0709)	(0.0633)
5km	0.162***						
	(0.0624)						
10km		0.0886					
		(0.108)					
20km			0.127				
			(0.135)				
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0267***	0.0260***	0.0270***				0.0256***
	(0.00611)	(0.00734)	(0.00838)				(0.00679)
5km				0.0256***			
				(0.00679)			
10km					0.0294***		
					(0.0106)		
20km						0.0314***	
						(0.00865)	
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km	0.0113	0.0176**	0.0209***	0.0240***	0.0277***	0.0202***	
	(0.00825)	(0.00799)	(0.00487)	(0.00438)	(0.00597)	(0.00427)	
5km							0.0240***
							(0.00438)
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	0.00492	0.00563	0.00650	0.00522	0.00534	0.00588	0.00522
	(0.00863)	(0.00821)	(0.00880)	(0.00804)	(0.00667)	(0.00722)	(0.00804)
(8) $\ln a_{it}^{\mathrm{POP}}$							
1km	-0.624	-0.628	-0.628	-0.660	-0.939*	-0.522	-0.660
	(0.472)	(0.517)	(0.546)	(0.490)	(0.497)	(0.504)	(0.490)
(9) τ_1	0.165***	0.166***	0.163***	0.173***	0.164***	0.166***	0.173***
	(0.0143)	(0.0174)	(0.0190)	(0.0150)	(0.0196)	(0.0117)	(0.0150)
(10) H. J p-value	0.952	0.972	0.974	0.928	0.938	0.878	0.928
(11) F	768.5	775.2	758.3	727.1	734	733.4	727.1
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

Table I.1: Regression results continued (Dependent variable: $\ln y_{it}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k_{it}^{D}$	0.284***	0.286***	0.284***	0.287***	0.280***	0.285***	0.280***
	(0.0249)	(0.0248)	(0.0274)	(0.0248)	(0.0280)	(0.0249)	(0.0264)
(2) $\ln k_{it}$	0.0932***	0.0923***	0.0913***	0.0931***	0.0907***	0.0921***	0.0903***
	(0.0115)	(0.0105)	(0.0121)	(0.0119)	(0.0117)	(0.0107)	(0.0117)
(3) $(\ln k_{it})^2$	-0.0822***	-0.0820***	-0.0817***	-0.0821***	-0.0817***	-0.0815***	-0.0809***
	(0.00864)	(0.00848)	(0.00867)	(0.00857)	(0.00861)	(0.00834)	(0.00861)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km	0.117*	0.124**	0.119*	0.118*	0.115*	0.116*	0.108*
	(0.0634)	(0.0599)	(0.0609)	(0.0608)	(0.0632)	(0.0654)	(0.0598)
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0251***	0.0240***	0.0209***	0.0274***	0.0177***	0.0275***	0.0271***
	(0.00780)	(0.00911)	(0.00498)	(0.00570)	(0.00635)	(0.00634)	(0.00530)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			0.0488***	0.0155	0.0283***	0.0218***	0.0217***
			(0.0159)	(0.0157)	(0.00899)	(0.00515)	(0.00695)
10km	-0.0245						
	(0.0165)						
20km		-0.105					
		(0.0783)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.00201	-0.00280				0.00736	0.0102*
	(0.0119)	(0.0129)				(0.00539)	(0.00619)
5km			0.0573***				
			(0.0199)				
10km				-0.0122			
				(0.0481)			
20km					0.137***		
					(0.0483)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.0295	
						(0.276)	
10km							0.666
							(0.547)
20km	-0.592	-0.491	-0.217	-0.806	0.232		
	(0.562)	(0.542)	(0.533)	(0.749)	(0.592)		
(9) τ_1	0.177***	0.190***	0.160***	0.178***	0.162***	0.189***	0.203***
	(0.0153)	(0.0173)	(0.0132)	(0.0159)	(0.0162)	(0.0213)	(0.0271)
(10) H. J p-value	0.943	0.944	0.875	0.934	0.846	0.920	0.935
(11) F	721.8	728.7	728.5	728.6	722.6	729.1	709.6
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table I.2: Regression results (Dependent variable: $\ln k^{D}_{it}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.381***	1.403***	1.404***	1.372***	1.378***	1.372***	1.372***
	(0.0616)	(0.0560)	(0.0571)	(0.0629)	(0.0583)	(0.0634)	(0.0629)
(2) $\ln k_{it}$	-0.0243	-0.0279	-0.0257	-0.0220	-0.0212	-0.0225	-0.0220
	(0.0660)	(0.0671)	(0.0671)	(0.0669)	(0.0673)	(0.0675)	(0.0669)
(3) $(\ln k_{it})^2$	0.226***	0.231***	0.230***	0.223***	0.224***	0.223***	0.223***
	(0.0184)	(0.0200)	(0.0216)	(0.0197)	(0.0207)	(0.0199)	(0.0197)
(4) $\ln a^{\mathrm{INV}}_{it}$							
1km				0.0138	0.0109	0.0144	0.0138
				(0.0426)	(0.0473)	(0.0418)	(0.0426)
5km	-0.0460						
	(0.0699)						
10km		-0.287***					
		(0.100)					
20km			-0.338**				
			(0.171)				
(5) $\ln a^{\mathrm{R\&D}}_{it}$							
1km	0.000665	0.00203	-0.000853				0.000705
	(0.00443)	(0.00617)	(0.00529)				(0.00478)
5km				0.000705			
				(0.00478)			
10km					-0.0150*		
					(0.00821)		
20km						0.00480	
						(0.0109)	
(6) $\ln a^{\mathrm{MNF_e}}_{it}$							
1km	-0.0106	0.00436	-0.00815	-0.0139	-0.0161	-0.0144	
	(0.0167)	(0.0151)	(0.0141)	(0.0147)	(0.0175)	(0.0146)	
5km							-0.0139
							(0.0147)
(7) $\ln a^{\mathrm{MNF_o}}_{it}$							
1km	0.00833	0.00795	0.00566	0.00814	0.0117	0.00741	0.00814
	(0.00995)	(0.00980)	(0.0111)	(0.00992)	(0.0112)	(0.00973)	(0.00992)
(8) $\ln a^{\mathrm{POP}}_{it}$							
20km	-0.539	-0.458	-0.471	-0.552	-0.688	-0.464	-0.552
	(1.231)	(1.061)	(1.023)	(1.229)	(1.247)	(1.179)	(1.229)
(9) τ_1	0.520***	0.555***	0.557***	0.514***	0.494***	0.520***	0.514***
	(0.0550)	(0.0441)	(0.0358)	(0.0504)	(0.0552)	(0.0471)	(0.0504)
(10) H. J p-value	0.253	0.243	0.251	0.255	0.253	0.254	0.255
(11) F	266	258.3	249.5	237.7	238	239.6	237.7
(12) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

Table I.2: Regression results continued (Dependent variable: $\ln k_{it}^{D}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.371***	1.377***	1.379***	1.381***	1.375***	1.366***	1.366***
	(0.0601)	(0.0652)	(0.0631)	(0.0595)	(0.0628)	(0.0614)	(0.0537)
(2) $\ln k_{it}$	-0.0228	-0.0232	-0.0212	-0.0205	-0.0228	-0.0233	-0.0227
	(0.0666)	(0.0661)	(0.0664)	(0.0648)	(0.0659)	(0.0652)	(0.0647)
(3) $(\ln k_{it})^2$	0.223***	0.224***	0.224***	0.225***	0.224***	0.223***	0.223***
	(0.0206)	(0.0194)	(0.0194)	(0.0197)	(0.0192)	(0.0185)	(0.0204)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km	0.0148	0.0176	0.0124	0.0220	0.0127	0.0117	0.0120
	(0.0421)	(0.0429)	(0.0443)	(0.0440)	(0.0434)	(0.0446)	(0.0491)
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.000500	9.21e-05	0.00515	0.0153**	0.000744	0.00240	0.00211
	(0.00483)	(0.00500)	(0.00487)	(0.00729)	(0.00518)	(0.00631)	(0.00686)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			-0.0372	-0.0761**	-0.0190*	-0.0183	-0.0154
			(0.0268)	(0.0386)	(0.00986)	(0.0151)	(0.0163)
10km	-0.0345						
	(0.0464)						
20km		-0.0559					
		(0.0614)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	0.00762	0.00926				0.0107	0.0103
	(0.00909)	(0.00726)				(0.00763)	(0.00854)
5km			-0.0337				
			(0.0369)				
10km				-0.152*			
				(0.0809)			
20km					0.0171		
					(0.0579)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.135	
						(0.450)	
10km							0.111
							(1.355)
20km	-0.566	-0.524	-0.954	-1.760	-0.529		
	(1.298)	(1.304)	(1.278)	(1.340)	(1.368)		
(9) τ_1	0.516***	0.522***	0.529***	0.553***	0.517***	0.529***	0.529***
	(0.0531)	(0.0574)	(0.0503)	(0.0535)	(0.0512)	(0.0384)	(0.0501)
(10) H. J p-value	0.256	0.252	0.258	0.258	0.256	0.258	0.252
(11) F	238.6	241.1	238.8	239.2	238.4	242.9	230.5
(12) #Obs	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table I.3: Regression results (Dependent variable: $\ln y_{it}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k_{it}^{D}$	0.355***	0.358***	0.351***	0.344***	0.344***	0.345***	0.344***
	(0.0290)	(0.0303)	(0.0312)	(0.0310)	(0.0285)	(0.0306)	(0.0310)
(2) $\ln k_{it}$	0.112***	0.108***	0.107***	0.114***	0.113***	0.111***	0.114***
	(0.0227)	(0.0211)	(0.0204)	(0.0228)	(0.0219)	(0.0231)	(0.0228)
(3) $(\ln k_{it})^2$	-0.176***	-0.175***	-0.175***	-0.178***	-0.177***	-0.177***	-0.178***
	(0.00650)	(0.00698)	(0.00731)	(0.00594)	(0.00608)	(0.00621)	(0.00594)
(4) $\ln a_{it}^{\text{INV}}$							
1km				0.200**	0.213**	0.212**	0.200**
				(0.0939)	(0.0957)	(0.104)	(0.0939)
5km	0.256***						
	(0.0789)						
10km		0.223					
		(0.149)					
20km			0.397**				
			(0.155)				
(5) $\ln a_{it}^{\text{R\&D}}$							
1km	0.0381***	0.0366***	0.0395***				0.0364***
	(0.0116)	(0.0124)	(0.0137)				(0.0127)
5km				0.0364***			
				(0.0127)			
10km					0.0399**		
					(0.0189)		
20km						0.0492***	
						(0.0154)	
(6) $\ln a_{it}^{\text{MNF}_e}$							
1km	-0.00656	-0.00148	0.00538	0.0132	0.0181	0.00757	
	(0.00776)	(0.00851)	(0.00928)	(0.00989)	(0.0126)	(0.0110)	
5km							0.0132
							(0.00989)
(7) $\ln a_{it}^{\text{MNF}_o}$							
1km	-0.00532	-0.00412	-0.00151	-0.00512	-0.00445	-0.00496	-0.00512
	(0.00882)	(0.00790)	(0.00931)	(0.00721)	(0.00467)	(0.00493)	(0.00721)
(8) $\ln a_{it}^{\text{POP}}$							
20km	0.112	0.0752	0.0741	0.0701	-0.340	0.361	0.0701
	(0.414)	(0.426)	(0.489)	(0.415)	(0.377)	(0.443)	(0.415)
(9) τ_1	0.158***	0.150***	0.137***	0.173***	0.159***	0.171***	0.173***
	(0.0376)	(0.0377)	(0.0338)	(0.0382)	(0.0366)	(0.0426)	(0.0382)
(10) H. J p-value	0.663	0.619	0.642	0.768	0.823	0.782	0.768
(11) F	588.2	593.8	568.5	557.6	564.3	563.9	557.6
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

Table I.3: Regression results continued (Dependent variable: $\ln y_{it}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k_{it}^{D}$	0.338***	0.344***	0.342***	0.345***	0.340***	0.343***	0.332***
	(0.0300)	(0.0299)	(0.0301)	(0.0313)	(0.0303)	(0.0289)	(0.0302)
(2) $\ln k_{it}$	0.114***	0.113***	0.112***	0.114***	0.111***	0.113***	0.109***
	(0.0225)	(0.0235)	(0.0233)	(0.0226)	(0.0235)	(0.0237)	(0.0229)
(3) $(\ln k_{it})^2$	-0.178***	-0.178***	-0.177***	-0.178***	-0.177***	-0.177***	-0.177***
	(0.00592)	(0.00566)	(0.00622)	(0.00596)	(0.00618)	(0.00549)	(0.00558)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.203**	0.212**	0.202**	0.199**	0.195**	0.198**	0.180**
	(0.0917)	(0.0878)	(0.0912)	(0.0893)	(0.0950)	(0.100)	(0.0886)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0356**	0.0341**	0.0285***	0.0355***	0.0252**	0.0367***	0.0353***
	(0.0147)	(0.0160)	(0.00994)	(0.0107)	(0.0120)	(0.0122)	(0.00869)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			0.0557***	0.0182	0.0276***	0.00800	0.0110
			(0.0202)	(0.0382)	(0.0101)	(0.0126)	(0.00690)
10km	-0.0728**						
	(0.0294)						
20km		-0.169					
		(0.106)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.0147**	-0.0128				-0.00360	0.00286
	(0.00740)	(0.00967)				(0.00616)	(0.00574)
5km			0.0767***				
			(0.0259)				
10km				0.00257			
				(0.0890)			
20km					0.168***		
					(0.0404)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.248	
						(0.331)	
10km							1.870***
							(0.672)
20km	0.155	0.286	0.803*	0.144	1.293**		
	(0.500)	(0.457)	(0.474)	(0.885)	(0.535)		
(9) τ_1	0.184***	0.199***	0.150***	0.169***	0.153***	0.177***	0.214***
	(0.0398)	(0.0322)	(0.0414)	(0.0358)	(0.0424)	(0.0299)	(0.0355)
(10) H. J p-value	0.776	0.654	0.850	0.729	0.842	0.809	0.767
(11) F	550.6	563.4	555.8	557.7	552.2	562.4	537.2
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table I.4: Regression results (Dependent variable: $\ln k_{it}^{D}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.743***	1.764***	1.743***	1.722***	1.730***	1.727***	1.722***
	(0.0748)	(0.0706)	(0.0784)	(0.0847)	(0.0799)	(0.0815)	(0.0847)
(2) $\ln k_{it}$	-0.0355	-0.0397	-0.0359	-0.0313	-0.0310	-0.0342	-0.0313
	(0.0635)	(0.0655)	(0.0659)	(0.0638)	(0.0647)	(0.0654)	(0.0638)
(3) $(\ln k_{it})^2$	0.266***	0.271***	0.266***	0.261***	0.262***	0.262***	0.261***
	(0.0226)	(0.0241)	(0.0243)	(0.0233)	(0.0250)	(0.0241)	(0.0233)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km				0.0800	0.0820	0.0849	0.0800
				(0.103)	(0.112)	(0.106)	(0.103)
5km	0.0170						
	(0.132)						
10km		-0.208					
		(0.150)					
20km			0.00578				
			(0.208)				
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0177**	0.0186*	0.0177**				0.0172*
	(0.00874)	(0.00989)	(0.00861)				(0.00896)
5km				0.0172*			
				(0.00896)			
10km					-0.00246		
					(0.0159)		
20km						0.0285	
						(0.0177)	
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km	-0.0451***	-0.0307*	-0.0441**	-0.0436**	-0.0442	-0.0464**	
	(0.0172)	(0.0159)	(0.0195)	(0.0219)	(0.0276)	(0.0218)	
5km							-0.0436**
							(0.0219)
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.0133	-0.0134	-0.0132	-0.0133	-0.00819	-0.0142	-0.0133
	(0.00886)	(0.00828)	(0.00945)	(0.00922)	(0.00928)	(0.0101)	(0.00922)
(8) $\ln a_{it}^{\mathrm{POP}}$							
20km	1.361	1.429	1.365	1.332	0.966	1.594*	1.332
	(1.039)	(0.896)	(1.030)	(1.050)	(1.078)	(0.895)	(1.050)
I2000	0.820***	0.850***	0.821***	0.814***	0.780***	0.821***	0.814***
	(0.0324)	(0.0283)	(0.0333)	(0.0285)	(0.0390)	(0.0280)	(0.0285)
(10) H. *J* p-value	0.358	0.348	0.363	0.363	0.352	0.363	0.363
(11) *F*	266	258.3	249.5	237.7	238	239.6	237.7
(12) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

Table I.4: Regression results continued (Dependent variable: $\ln k_{it}^{D}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.719***	1.734***	1.720***	1.724***	1.717***	1.728***	1.711***
	(0.0840)	(0.0892)	(0.0827)	(0.0786)	(0.0817)	(0.0811)	(0.0761)
(2) $\ln k_{it}$	-0.0339	-0.03 49	-0.0293	-0.0278	-0.0305	-0.0333	-0.0337
	(0.0627)	(0.0624)	(0.0647)	(0.0628)	(0.0642)	(0.0618)	(0.0605)
(3) $(\ln k_{it})^2$	0.261***	0.264***	0.260***	0.261***	0.260***	0.262***	0.259***
	(0.0234)	(0.0224)	(0.0235)	(0.0237)	(0.0234)	(0.0232)	(0.0238)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km	0.0832	0.0902	0.0800	0.0935	0.0811	0.0756	0.0647
	(0.103)	(0.102)	(0.106)	(0.0974)	(0.102)	(0.114)	(0.108)
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0165*	0.0156*	0.0187*	0.0348**	0.0156	0.0143	0.0125
	(0.00931)	(0.00926)	(0.0107)	(0.0154)	(0.0104)	(0.00888)	(0.00974)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			-0.0511	-0.114	-0.0338*	-0.0512*	-0.0425*
			(0.0434)	(0.0779)	(0.0192)	(0.0272)	(0.0221)
10km	-0.107*						
	(0.0572)						
20km		-0.148**					
		(0.0729)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.0147*	-0.00892				-0.0138	-0.0108
	(0.00890)	(0.00783)				(0.00843)	(0.0109)
5km			-0.0346				
			(0.0540)				
10km				-0.219			
				(0.156)			
20km					-0.00425		
					(0.0639)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.494	
						(0.529)	
10km							1.445
							(1.530)
20km	1.287	1.390	1.178	-0.112	1.461		
	(1.112)	(1.129)	(1.043)	(1.348)	(1.194)		
(9) τ_1	0.819***	0.834***	0.817***	0.856***	0.807***	0.795***	0.809***
	(0.0296)	(0.0348)	(0.0330)	(0.0489)	(0.0301)	(0.0177)	(0.0318)
(10) H. *J* p-value	0.361	0.359	0.361	0.359	0.362	0.356	0.348
(11) *F*	238.6	241.1	238.8	239.2	238.4	242.9	230.5
(12) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

J First-stage results for the IV regressions based on indirect collaborators in different firms

Tables J.1 and J.2 show the results form the first-stage regressions corresponding to Tables F.1 and F.2, respectively.

Table J.1: First-stage regression results for (5.1) under alternative IVs

	Quality				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) k_{it}^{D,IV_3}	0.239***	0.267***			0.251***	0.313***		
	(0.0254)	(0.0110)			(0.0203)	(0.0111)		
(2) k_{it}^{D,IV_4}	0.0592**		0.226***		0.106***		0.314***	
	(0.0235)		(0.0215)		(0.0279)		(0.0150)	
(3) k_{it}^{D,IV_5}	-0.0131			0.163***	0.0282			0.277***
	(0.0493)			(0.0417)	(0.0338)			(0.0267)
(4) $\ln k_{it}$	0.130***	0.130***	0.134***	0.136***	0.158***	0.159***	0.167***	0.170***
	(0.0331)	(0.0330)	(0.0356)	(0.0369)	(0.0126)	(0.0128)	(0.0163)	(0.0163)
(5) $(\ln k_{it})^2$	-0.0513***	-0.0513***	-0.0525***	-0.0532***	-0.0878***	-0.0882***	-0.0891***	-0.0897***
	(0.0117)	(0.0116)	(0.0131)	(0.0136)	(0.00845)	(0.00836)	(0.0102)	(0.0107)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.419***	0.419***	0.428***	0.431***	0.553***	0.561***	0.573***	0.581***
	(0.0777)	(0.0770)	(0.0779)	(0.0784)	(0.104)	(0.105)	(0.108)	(0.112)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00353	0.00387	0.00520	0.00810	0.0145	0.0153	0.0186	0.0242
	(0.0100)	(0.0102)	(0.0105)	(0.0110)	(0.0150)	(0.0157)	(0.0162)	(0.0175)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0640**	-0.0640**	-0.0652**	-0.0652**	-0.110***	-0.110***	-0.117***	-0.120***
	(0.0270)	(0.0265)	(0.0295)	(0.0302)	(0.0201)	(0.0177)	(0.0232)	(0.0259)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0191	0.0198	0.0182	0.0196	0.00403	0.00514	0.00312	0.00233
	(0.0245)	(0.0236)	(0.0255)	(0.0251)	(0.0283)	(0.0260)	(0.0304)	(0.0296)
(10) $\ln a_{it}^{\mathrm{POP}}$	0.652	0.708	0.674	0.855	2.930**	3.063***	2.971**	3.219**
	(1.281)	(1.199)	(1.326)	(1.297)	(1.249)	(1.136)	(1.326)	(1.300)
(11) τ_1	0.318***	0.324***	0.343***	0.369***	0.498***	0.524***	0.549***	0.587***
	(0.0383)	(0.0314)	(0.0396)	(0.0430)	(0.0276)	(0.0300)	(0.0299)	(0.0332)
(12) R^2	0.182	0.181	0.172	0.165	0.194	0.193	0.185	0.177
(13) F	268.1	593.7	110	15.33	494	799	437.2	108.1
(14) p-value	0	0	0	0.000129	0	0	0	0
(15) #Obs.	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

Table J.2: First-stage regression results for (5.10) under alternatie IVs

	Quality				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) $\ln \Delta n_{it}^{\mathrm{IV}_3}$	0.247***	0.233***			0.247***	0.233***		
	(0.00775)	(0.00780)			(0.00775)	(0.00780)		
(2) $\ln \Delta n_{it}^{\mathrm{IV}_4}$	-0.0345**		0.219***		-0.0345**		0.219***	
	(0.0153)		(0.00834)		(0.0153)		(0.00834)	
(3) $\ln \Delta n_{it}^{\mathrm{IV}_5}$	0.0209			0.200***	0.0209			0.200***
	(0.0160)			(0.00837)	(0.0160)			(0.00837)
(4) $\ln k_{it}$	0.0655**	0.0657**	0.0746***	0.0811***	0.0655**	0.0657**	0.0746***	0.0811***
	(0.0268)	(0.0267)	(0.0285)	(0.0296)	(0.0268)	(0.0267)	(0.0285)	(0.0296)
(5) $(\ln k_{it})^2$	-0.165***	-0.165***	-0.171***	-0.175***	-0.165***	-0.165***	-0.171***	-0.175***
	(0.00833)	(0.00829)	(0.00939)	(0.00999)	(0.00833)	(0.00829)	(0.00939)	(0.00999)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.136***	0.135***	0.134***	0.156**	0.136***	0.135***	0.134***	0.156**
	(0.0433)	(0.0430)	(0.0511)	(0.0615)	(0.0433)	(0.0430)	(0.0511)	(0.0615)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00555	0.00583	0.00413	0.000824	0.00555	0.00583	0.00413	0.000824
	(0.00561)	(0.00546)	(0.00584)	(0.00599)	(0.00561)	(0.00546)	(0.00584)	(0.00599)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	0.0210*	0.0217*	0.0233	0.0126	0.0210*	0.0217*	0.0233	0.0126
	(0.0127)	(0.0123)	(0.0144)	(0.0189)	(0.0127)	(0.0123)	(0.0144)	(0.0189)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0106	0.0104	0.0133	0.0164	0.0106	0.0104	0.0133	0.0164
	(0.00826)	(0.00797)	(0.00856)	(0.0104)	(0.00826)	(0.00797)	(0.00856)	(0.0104)
(10) $\ln a_{it}^{\mathrm{POP}}$	0.403	0.409	0.568**	0.671**	0.403	0.409	0.568**	0.671**
	(0.278)	(0.280)	(0.281)	(0.304)	(0.278)	(0.280)	(0.281)	(0.304)
(11) τ_1	-0.177***	-0.175***	-0.189***	-0.202***	-0.177***	-0.175***	-0.189***	-0.202***
	(0.0140)	(0.0132)	(0.0128)	(0.0142)	(0.0140)	(0.0132)	(0.0128)	(0.0142)
(12) R^2	0.329	0.329	0.305	0.284	0.329	0.329	0.305	0.284
(13) F	470	892	692	571	470	892	692	571
(14) p-value	0	0	0	0	0	0	0	0
(15) #Obs.	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1