Do Individuals or Firms Matter More? The Case of Patent Generation

Tong Liu
Peking University
liutongccer@pku.edu.cn
(86) 10-62751475

Yifei Mao
Cornell University
ym355@cornell.edu
607-255-8140

Xuan Tian
Indiana University and Tsinghua University
tianx@indiana.edu
812-855-3420

This version: February, 2016

Do individuals or firms matter more? The case of patent generation

Abstract

This paper studies the relative importance of firm organizational capital and individual inventor human capital in promoting a firm's innovation output. We decompose the variation in innovation output into firm- and inventor-specific components. Inventor human capital is about 13 times more important than firm organizational capital in explaining a firm's innovation performance in terms of patent counts and citations, while inventor human capital is only about two times more important when explaining the firm's innovation style in terms of patent exploratory and exploitive scores. In the cross section, inventors contribute more to innovation output when they are better networked, in firms with higher inventor mobility, in high-tech industry firms, and in publicly traded firms. This paper highlights the importance of individual inventors in enhancing firm innovation and sheds new light on the theory of the firm.

Key words: Organizational capital, Human capital, Firms, Inventors, Innovation

JEL number: G30, G32, O32

1. Introduction

Since Coase (1937), there has been a longstanding debate on what constitutes a firm. On the one hand, in the Hart-Moore framework, nonhuman assets are the glue that brings a firm together (Hart (1995)). On the other hand, Zingales (2000) argues that "human capital is emerging as the most crucial asset" in today's world. It is still unclear to researchers, however, what role a firm's organizational capital and its human capital play in explaining a firm's long-run success. Is it the growth engine of a firm embedded in its human capital or non-human capital such as organizational structure, culture, access to resources and others that drive the firm's growth and success? How important is the human capital in determining a firm's growth strategy and outcome? This paper attempts to answer these questions.

To this end, we focus on the engine to a firm's competitive advantage and long-run success, namely, innovation. Specifically, we isolate the contribution of the firm and its inventors to firm-level innovation. Innovation is difficult to achieve, requiring investment in human capital and significant tolerance of early failure (Manso (2011) and Tian and Wang (2014)). While there has been many studies investigating various market and firm characteristics that contribute to innovation, little is known about the separate roles played by inventors and firms in driving innovation. This topic is of particular interest to corporate decision makers and policy makers in their attempts to promote innovation. This paper provides the first empirical study that quantifies how inventors and firms matter in innovation.

Shedding light on the sources of innovation performance and "styles" improves our understanding of what constitutes a firm. While both investment performance and "styles" could vary in firms, we investigate two corresponding dimensions of innovation. The first dimension is innovation performance, measured by both the quantity and quality of innovation output. The second dimension is innovation "styles". There are two main styles in innovative investment documented in the management literature: exploratory innovation and exploitative innovation (Gao et al. (2015)). Exploratory innovation is radical innovation that requires knowledge that is outside the existing knowledge domain. In contrast, exploitative innovation is incremental innovation that builds on existing knowledge and improves existing skills, processes, and structures.

A major challenge this paper has is to distinguish between the contribution from organizational capital and that from human capital. Innovation provides a unique setting that is

clean and rich when dealing with this empirical challenge, given that we could track individual inventors' patent filings and the citations received by their patents. On top of that, we are able to observe inventors move from one firm to another, using the Harvard Business School (HBS) Inventor Patent Citation database. Intuitively, if an individual's output does not change whichever firm it moves to, then the individual is the sole driver of his output. If an individual's output changes when he moves from one firm to another firm, then we can largely attribute the change in the output to the second company.

We employ two methods to isolate the role of the firm and the inventor. The first method is to examine a panel of inventors that have changed their affiliated firms and to include inventor, firm, and year fixed effects in the specification. We refer to this approach as the mover dummy variable (henceforth MDV) method, which has been used commonly in existing literature (e.g., Bertrand and Schoar (2003), Graham, Li, and Qiu (2012)). However, because the MDV approach is limited to the movers only, which is only 16% of all the inventors in our sample, we use an alternative method that includes both movers and stayers in the sample, as long as the stayers are in firms that employee at least one mover. This method is developed by Abowd, Kramarz, and Margolis (1999) (henceforth AKM) and later refined by Abowd, Creecy, and Kramarz (2002). The AKM method extends the rather small sample of movers to the connectedness sample, which includes 98.4% of all the inventors.

The key results and their economic implications from our paper are similar regardless of the method we use. The inventor fixed effects are consistently more important than firm fixed effect in explaining innovation output, no matter it is measured by patent counts, citations, exploratory score, or exploitative score. Specifically, our estimates suggest that inventor human capital is about 13 times more important than firm organizational capital in explaining the firm's innovation performance (patent counts and citations), while inventor human capital is only about two times, compared to firm organizational capital, more important when explaining the firm's innovation style (patent exploratory and exploitive scores). The results suggest that while human capital is crucial in determining innovation performance, its effect on a firm's innovation style is much more moderated. In other words, as inventors switch firms, their performance in terms of high quality patent filings is inherently determined by the inventors' own abilities, but the radicalness and originality of innovation they pursue would be largely affected by their affiliated firms' organizational environment.

Furthermore, we find that the inventor fixed effect estimates demonstrate significant heterogeneity in the cross section. To understand heterogeneous effects, we explore how they vary in subsamples partitioned in a number of dimensions. First, we examine inventor networks. The degree of an inventor's centrality is determined by the number of coauthor relationships she has. We postulate that an inventor at the center of coauthor relationships files more high quality patents. However, these inventors do not necessarily pursue more or less radical innovation, because the type of innovation is largely affected and by the firm. The results are consistent with our conjectures. We find that, compared to an average inventor in terms of network centrality, center inventors contribute significantly to patent counts and citations, but they do not contribute significantly to patent exploratory score or exploitative score.

Second, we investigate how the mobility of inventors in a firm alters our main results. High inventor turnovers in a firm could be a result of two reasons. First, the inventors' contribution to innovation in the firm is not as important as the firm's organizational capital, and therefore these inventors are replaceable. Second, firms with high inventor mobility could have more talented inventors and talented inventors tend to have higher mobility than other inventors. We find that the second effect dominates. Inventors in high mobility firms appear to be more important in determining innovation output compared to inventors in low mobility firms.

Third, we explore industry heterogeneity and focus on high-tech industries for which innovation is particularly important. It is likely that in these industries inventor human capital plays a dominating role than the firm. We find that, in drug, chemical, computer and electrical industries that are typically considered high-tech industries, inventors are more crucial in driving innovation output than those in other industries. The findings support our hypothesis.

Finally, we compare inventor contribution in public firms versus private firms. Public firms, being more transparent and subject to higher short-term pressures from various market players, tend to attract more homogeneous talented inventors. Private firms, on the other hand, would bear more heterogeneous inventors with different backgrounds and skill sets. Therefore, it is likely that we would observe a higher inventors' contribution to innovation output in public firms than in private firms. Our findings are consistent with this conjecture.

The primary contribution of our paper is threefold. First, our paper contributes to the literature on the economics of organization. This literature proposes different hypotheses for the existence of the firm and distinguishes physical from human capital (Coase (1988), Klein (1988),

Williamson and Winter (1993)). As an empirical study, Kaplan et al. (2009) examine startup companies and suggest the business (nonhuman capital) is more important than the management team (human capital). Interestingly, our results highlight the importance of human capital in established firms, suggesting the differences in the nature of startups and established companies, which supplements Kaplan et al. (2009).

Second, our paper contributes to the innovation literature by documenting the importance of inventor human capital in explaining innovation performance and style. To the best of our knowledge, this paper is the first one that isolates the contribution of inventor human capital from firm organizational capital in determining innovation output. Prior studies have examined multiple determinants of innovation, including legal environment (Acharya et al. (2014)), banking competition (Cornaggia et al. (2015)), financial market development ((Hsu et al. (2014)) institutional ownership (Aghion et al. (2013)), product market competition (Aghion et al. (2005)), etc. However, these studies examine innovation at the firm level, and only study extensive margins that affect firm-level innovation. In this paper, we delve deeper at the intensive margin and decompose innovation drivers into human capital-related and firm organization capital-related components, which allows us to further understand relative importance of these two components in promoting firm innovation output.

Third, this paper also contributes to the expanding literature that attempts to separate human capital and organizational capital. Abowd, Kramarz, and Margolis (1999) find that individual effects are more important than firm effects in explaining wage variations in France. Bertrand and Schoar (2003) and Graham, Li, and Qiu (2011) show that manager fixed effects explain a significant extent of firm policy heterogeneity and are related to management styles, and managers with higher performance fixed effects receive higher compensation. Ewens and Rhodes-Kropf (2013) find that venture capitalists have repeatable skills and VC partner's human capital is more important than VC firm's organizational capital in explaining performance. Berk, Binsbergen, and Liu (2014) stress the role of mutual fund firms in efficiently allocating capital to their managers. Existing studies are either unable to capture individuals' output in the setting of workers or firm/fund managers, or have to infer individual output through indirect ways as in the VC partner setting. In this paper, we are able to directly track individual inventor output by using patents filed by each one of them. This unique feature of inventor/patent data provides a clean

setting to test our conjectures and enables us to delve deeply and examine individual inventors' performance and style.

The rest of the paper is organized as follows. Section 2 describes data and variable constructions. Section 3 reports the main results. Section 4 presents subsample results that explore the heterogeneity of our main findings. Section 5 concludes.

2. Data

2.1 Sample construction

We begin with the latest version of the Harvard Business School (HBS) patent and inventor database available at http://dvn.iq.harvard.edu/dvn/dv/patent. [1] The HBS patent and inventor database provides information for both inventors (the individuals who receive credit for producing the patent) and assignees (the entity that owns the patents, which could be a government, a firm, an organization, or an individual). For the purpose of our study, we need to track the employer of the inventor as he moves around from one firm to another. Since patent database does not contain information on employment of the inventor, we assume that the employer of the inventor is the company to which the patent (filed by the inventor) is assigned. There is a clear identification of the employer if a patent is only assigned to one assignee. However, when a patent is assigned to multiple assignees, the HBS patent and inventor database only reports the primary assignee of the patent. This issue confounds the identification of the employer of then inventor. To overcome this problem, we match the HBS patent and inventor database with the National Bureau of Economics Research (NBER) Patent Citation database that contains precise patent and assignee information. [2] We discuss in more detail regarding how to pin down the employer for each inventor when there are multiple assignees in Appendix A.

Next, to obtain time-varying firm characteristics, we merge the inventor-year patent sample with the firm-level annual accounting variables obtained from Compustat. We require all firms to have non-missing financial records across our sample period. Finally, we omit observations before 1970 (i.e., 345 observations), which is only a small portion of our final

---

[1] See Li et al. (2014) and Singh and Fleming (2010) for more details about the HBS patent and inventor database.

[2] This database is available for downloading at https://sites.google.com/site/patentdataproject/Home/downloads. See Hall, Jaffe, and Trajtenberg (2001) for more details about the NBER patent citation database.

sample). Our final sample consists of 204,678 inventors (1,250,041 inventor-year observations), that have worked for 5,722 firms from 1970 to 2003.[3]

2.2 Variable measurement

2.2.1 Measuring Innovation

We construct two sets of patent-based metrics to gauge an inventor's innovation outcome. The first set is to measure innovation performance and the second set is to capture their innovation style. Following the innovation literature, one measure in the first set is the total number of patents filed and eventually granted in a given year by an inventor, which captures the quantity of innovation. We use the application year instead of the grant year to determine an inventor's innovation output because the patent application year is closer to the actual time when innovation activities take place (Griliches, Pakes, and Hall, 1988; Griliches, 1990). Although the intuition is straightforward and it is easy to implement, a simple measure of patent counts hardly distinguishes breakthrough innovations from incremental technological discoveries (Trajtenberg, 1990). Hence, we construct another metric of innovation output, the total number of non-self citations each patent receives in subsequent years. We use this measure to capture the quality (or the impact) of innovation.

Nevertheless, both innovation measures suffer from severe truncation problems. Because in our matched sample we only observe patents that are eventually granted by the end of 2006, patents filed in the last few years may still be under review and not granted by 2006 (this truncation problem is mainly caused by using the NBER database to match). To adjust the truncation bias in patent counts, we calculate each inventor's patent of a given year in the HBS database, which contains patents granted through 2010. To the extent that the patent application outcomes have been announced by 2010 for the patents filed by 2006, this approach greatly alleviates the patent truncation concern. Similarly, patents tend to receive citations over a long period after its grant date, but we observe at best the citations received up to 2010.[4] To deal with this truncation bias, we correct the citation data by using the "weight factors" following Hall, Jaffe, and Trajtenberg (2001, 2005) and estimating the shape of the citation-lag distribution.

---

[3] The time span of our sample ends in 2003 because we take the innovation metrics three years ahead in our analysis in order to capture the long-term nature of innovation activity.

[4] As the HBS patent database contains citation information up to 2010, our best observing window closes in 2010. However, this does not help too much in adjustment for the citation truncation problem because patents may keep receiving citations over a long period of time beyond 2010.

Consistent with the innovation literature, the distribution of patent grants in our final sample is right skewed, with its median at zero. Due to the right skewness of patent counts and citations per patent, we use the natural logarithm of one plus patent counts (*LnPatent*) and the natural logarithm of one plus the number of citations per patent (*LnCitation*) as the main innovation metrics to measure innovation performance in our analysis. We also winsorize all our dependent variables at the 99th percentile.

The second set of metrics includes *Exploit* and *Explore* indices, which reflect an inventor's innovation style. We follow existing literature (e.g., Sorensen and Stuart (2000), Katila and Ahuja (2002), etc.) to category an inventor's patent activity into exploratory innovation and exploitative innovation. The basic idea is that inventors concentrating on their existing knowledge are expected to produce more exploitative patents while inventors exploring new ideas are expected to create more exploratory patents. We define an inventor's existing knowledge as his previous patent portfolio and the set of patents that has been cited by her own patents over the past five years. Then proxies are constructed such that a patent is classified exploitative if at least 60% of its citations are based on existing knowledge, and a patent is classified exploratory if at least 60% of its citations are based on new knowledge.[5] We then set *Exploit* equal to the ratio of the number of exploitative patents for a given year to the total number of patents filed by the inventor in the same year. In a similar way, we define *Explore* by the ratio of the number of exploratory patents for a given year to the total number of patents filed by the inventor in the same year.

Note that the patent databases used in our study are unlikely to be affected by survivorship bias. As long as a patent application is eventually granted, it is attributed to the applying firm at the time of application even if the firm later gets acquired or goes bankrupt. Moreover, because patent citations are attributed to the patent rather than the assignee, the patent granted to a firm that later gets acquired or goes bankrupt can still keep receiving citations long after the firm disappears. For firm characteristics, we compute all variables for firm *i* over its fiscal year *t*. Our control variables include firm size (the natural logarithm of book value assets), firm age (the natural logarithm of a firm's age since its IPO year), profitability (ROA), investments in intangible assets (R&D expenditures over total assets), asset tangibility (net PPE scaled by total assets), leverage, capital expenditures, growth opportunities (Tobin's Q), financial

---

[5] We also use 80% as a cutoff point. The results are robust to that.

constraints (the Kaplan and Zingales (1997) five-variable KZ index), and industry concentration (the Herfindahl index based on sales). Aghion et al. (2005) points out non-linear effects of product market competition on innovation output. Hence, we include the squared Herfindahl index in our regressions. We provide detailed variable definitions in the Appendix B.

2.3 Summary statistics

Table 1 presents information on the movers and stayers in our sample. Panel A shows that during the sample period, 15.91% of (or 32,561) inventors are movers who work in more than one firm in the sample, while the rest (84.09%) are non-movers who work in the same sample firm throughout our sample period. Panel B provides information on the proportion of firms that have a given number of movers during the sample period. 24.68% of the sample firms do not have any movers while the remaining 75.32% of (4,310) firms have at least one mover. According to the AKM method, we are able to identify the fixed effects of inventors who work in these 4,310 firms regardless of whether they are movers or stayers, which constitutes the connectedness sample. In the robustness check, we perform the MDV method on the mobility sample, which comprises of 32,561 movers as well as 4,310 firms at which these movers are employed.

[Insert Table 1 Here]

2.2.2 Measuring control variables

Following the innovation literature, we control for a vector of inventor and firm characteristics that may affect innovation performance. For inventor time-varying characteristics, we create proxies for inventors' prior innovation experience. Two variables, *LnExpnum* and *LnExpcit*, are defined as the logarithm of one plus the average metrics (adjusted patent count and patent citation, respectively) over the past three years. We use a three-year rolling window because recent experience is a good indicator that the inventor is an active participant in innovation (Chemmanur, Ertugrul, and Krishnan, 2015). Construction of this variable requires information on the past three years' invention experience, and hence we exclude the first three-year observations for all inventors because their prior innovation experience is missing. We also exclude the inventors whose moving happens in the first three years in our sample period so that we can keep the mobility structure intact.

[Insert Table 2 Here]

To minimize the effect of outliers, we winsorize all control variables at the 1st and 99th percentiles. Table 2 provides summary statistics of the variables in both the full sample and the connectedness sample, in order to examine whether the connectedness sample is representative of the full sample (Brav et al., 2005).[6] Panel A summarizes the representativeness of these variables for inventors. On average, an inventor in the sample has 0.97 granted patents per year and each patent receives 6.90 citations. In the connectedness sample, an inventor has a similar number of patents granted per year, 0.97 and each patent receives 6.91 citations. The other variables that measure inventors' prior innovation experience are also very close in both the full sample and the connectedness sample.

Table 2 Panel B summarizes the representativeness of these variables for firms. In the full sample, an average firm has book value assets of $7.12 billion, R&D-to-assets ratio of 5.9%, ROA of 11.5%, PPE-to-assets ratio of 29.8%, leverage of 22.2%, capital expenditure ratio of 6.8%, Tobin's Q of 1.9, and is 21.9 years old since its IPO date. In the connectedness sample, these statistics are quite close: an average firm has book value assets of $7.57 billion, R&D-to-assets ratio of 6.0%, ROA of 12.1%, PPE-to-assets ratio of 29.9%, leverage of 22.0%, capital expenditure ratio of 6.9%, Tobin's Q of 1.9, and is 22.4 years old since its IPO date. Overall, these comparisons allow us to infer that our connectedness sample is representative of the universe of inventors and firms in the full sample.

3. Main results

3.1 Empirical methods

Our empirical tests relate inventor and firm characteristics in the current year to the metrics of innovation performance three years ahead in view of the long-term nature of innovation process. Consider the following linear model of inventors' innovation $Y_{ij(t+3)}$:

$$Y_{ij(t+3)} = \beta_1 X_{it} + \beta_2 Z_{jt} + \phi_i + \theta_j + \mu_t + \epsilon_{ijt} \tag{1}$$

where i denotes inventor, j denotes firm, and t denotes year when the innovation activity occurs. In above equation, X_{it} and Z_{jt} include time-varying controls for inventor and firm. μ_t captures

---

[6] When using *Exploit* and *Explore* indices as dependent variables in the baseline regression, the sample size is different since we assign a missing value for an inventor of a year in the case that no patent was filed by him or her. We provide summary statistics of *Exploit* and *Explore* for both the full sample and the connectedness sample in the online appendix.

the year fixed effects. Our focus is to retrieve both inventor and firm fixed effects ϕ_i and θ_j utilizing movements of inventors across firms.

We use a method first proposed by Abowd, Kramarz and Margolis (1999) (hereafter referred to as AKM method) and later refined by Abowd, Creecy and Kramarz (2002). The AKM method allows us to separate firm and inventor fixed effects through connectedness sample, which includes not only moving inventors but also non-movers who work in firms that have hired at least one mover. To define a connectedness sample, we use graph theory to determine groups of inventors and firms that are connected. Detailed procedures are as follow: Start with an arbitrary inventor and track all firms where he or she has ever worked. Then we include all inventors whoever work in the firms mentioned above into our connectedness sample and continue tracking all firms for which these inventors have ever worked. Repeat the procedure until all data are exhausted.[7] Abowd, Creecy, and Kramarz (2002) show that connections make the estimation of inventor and firm fixed effects for each connected group relative to a within-group benchmark computationally feasible. To make inventor and firm fixed effects directly comparable across groups, we follow the normalization procedure suggested by Cornelissen (2008): First, we normalize the mean firm fixed effects for each group to zero and add the group mean firm fixed effects to inventor fixed effects; Second, we subtract the grand mean of inventor fixed effects from each inventor fixed effect and then add this grand mean inventor fixed effect to the intercept.

An analogous method (i.e., the MDV method) proposed by Bertrand and Schoar (2003) employs the mobility sample consisting of only movers and firms for which they work to separate firm fixed effects from individual fixed effects, using the LSDV approach. One disadvantage of the MDV method compared to the AKM method is a potential sample selection bias resulting from the restriction of sample to only movers who could be different from non-movers. Besides what we mentioned above, there are some other important benefits of adopting the AKM framework. First, the AKM method uses information of both movers and non-movers, which gives us larger sample size and higher statistical power. Secondly, it can significant reduce

---

[7] In most of our analyses, we use Cornelissen's (2008) Stata command "felsdvreg" to implement the AKM method and estimate both inventor and firm fixed effects. This command facilitates the estimation of a linear model with two high-dimensional fixed effects (i.e., inventor and firm fixed effects) by using a memory-saving decomposition of the design matrix. It also provides useful summary statistics when taking care of identification issues. In some situations with tremendous data size, we switch to the Stata command "reghdfe" proposed by Correia (2014), which is more efficient when dealing with data that requires large memory.

the computational work in terms of the large dataset used in our study. Nonetheless, we still conduct the MDV approach in our analysis as a robustness check

We now provide a detailed discussion on how the AKM method separately identifies inventor and firm fixed effects using the connectedness sample, because it's useful to understand the basic futures of the AKM to better appreciate its advantages and realize its limitations. Define the variable D_{ikt} as a dummy that is equal to one if inventor i works at firm k at time t and zero otherwise. Then we can rewrite equation (1) as:

$$Y_{ij(t+3)} = \beta_1 X_{it} + \beta_2 Z_{jt} + \phi_i + \sum_{k=1}^{J} D_{ikt}\theta_k + \mu_t + \epsilon_{ijt}. \tag{2}$$

In the first step, the AKM approach sweeps out the inventor fixed effects by averaging over all inventor i's innovation performances to obtain:

$$\bar{Y}_i = \beta_1 \bar{X}_i + \beta_2 \bar{Z}_i + \phi_i + \sum_{k=1}^{J} \bar{D}_{ik}\theta_k + \bar{\mu}_t + \bar{\epsilon}_i. \tag{3}$$

Here $\bar{Y}_i$ is inventor i's average innovation performance across the full sample period. Then we begin to demean (2) with (3) in order to get:

$$Y_{ij(t+3)} - \bar{Y}_i = \beta_1(X_{it} - \bar{X}_i) + \beta_2(Z_{jt} - \bar{Z}_i) + \sum_{k=1}^{J}(D_{ikt} - \bar{D}_{ik})\theta_k + (\mu_t - \bar{\mu}_t) + (\epsilon_{ijt} - \bar{\epsilon}_i). \tag{4}$$

Through demeaning process the inventor fixed effects have been removed. Now it's clear that we are able to exploit movers' information to identify firm fixed effects since $D_{ikt} - \bar{D}_{ik} \neq 0$ for a mover, which can be estimated by the LSDV method. Finally, using the estimates in above regression, we can recover the inventor fixed effects following equation:

$$\hat{\phi}_i = \bar{Y}_i - \hat{\beta}_1 \bar{X}_i - \hat{\beta}_2 \bar{Z}_i - \sum_{k=1}^{J} \bar{D}_{ik}\hat{\theta}_k \tag{5}$$

and here $\bar{\mu}_t$ is often treated as the benchmark in estimating time effects and thereby assumed to be zero.

As Abowd et al. (2004) and Andrews et al. (2008) note, an estimation bias may emerge when inventor mobility is limited, which could lead to imprecise estimation of inventor and firm fixed effects. Consequently, we need to be cautious when interpreting the results in both the MDV and AKM methods. However, this issue is not severe in our study as our sample contains about 16% movers. This proportion is much higher than that in previous literature (e.g. Graham, Li and Qiu (2012)). Another property of the AKM estimator is that fixed effects estimates themselves have properties that are similar to other estimators. As shown by Wooldridge (2010), the estimates of the time-varying variable coefficients are unbiased and consistent, while the fixed effects estimates are only unbiased.

3.2 Baseline results

In this section, we analyze how unobserved inventor and firm time invariant characteristics affect inventors' innovation outcomes using the AKM method. The AKM method uses a connectedness sample that eliminates firms that do not have a mover during our sample period. Based on this procedure, the connectedness sample in the AKM estimates for innovation performance has 201,461 inventors (32,561 movers), 4,310 firms, and 1,239,614 inventor-year observations, which accounts for 98% of all inventors, 75% of all firms, and 99% of all observation units.

We follow the prior literature to select the observable characteristics of inventors and firms that may affect an inventor's future innovation performance (e.g., He and Tian, 2013; Tian and Wang, 2014; Seru, 2014; Cornaggia et al., 2015.). Specifically, in our full fixed effects model we regress the proxy of inventors' innovation performance on both firm time-varying variables, such as firm size, age, profitability, intangible asset, and on inventor time-varying variables, such as prior experience of inventor. Additionally, we include year fixed effects to capture the impact of economic conditions.

Table 3 reports the results of estimating equation (1) using the AKM method in the connectedness sample. We suppress all coefficient estimates and focus on the relative economic importance of time-invariant inventor and firm characteristics. Following Graham, Li and Qiu (2012) and Ewens and Rhodes-Kropf (2015), we use $\frac{cov(Y, Inventor\ FE)}{var(Y)}$ to examine the contribution of inventor fixed effects to the total variation in inventors' innovation outcome. $\frac{cov(Y, Inventor\ FE)}{var(Y)}$ reports the covariance of the dependent variable with inventor fixed effects, scaled by the dependent variable's variance. These normalized covariance term represents the fractions of total variations attributable to particular factors, which can effectively capture the relative importance of different fixed effects in explaining the dependent variable for a given regression model. In addition, we report the incremental change in adjusted-R square across three different model specifications in Table 3: the first specification in which we include all control variables and year effects; the second specification in which we include all control variables, year effects, and firm fixed effects; the third specification in which we include all control variables, year effects, and inventor fixed effects; the final specification in which we include all control variables and all fixed effects using the AKM method. Adjusted R-squares are

adopted in this case because the number of explanatory variables changes across models.

[Insert Table 3 Here]

For patent counts in column 1, inventor fixed effects account for 67.8% of the total variation while firm fixed effects contribute to 5.5% of the total variation (the left portion is attributable to all other controls). The relative importance of inventor fixed effects compared to firm fixed effects is measured by the ratio of the contribution of these two fixed effects, which is 13 times in column 1. For patent citations in Column 2, 90.4% of the total variation corresponds to inventor fixed effects while 6.33% of the total variation corresponds to firm fixed effects. The relative importance between inventor and firm fixed effects amounts to about 14 times. Overall, for patent counts and citations, the stark differences in explanatory power between inventor and firm fixed effects reflect that innovation performance is largely driven by inventor fixed effects. These results show the important role of inventors' inherent ability or time invariant characteristics, compared to firms' time invariant characteristics, in shaping innovation outcomes.

In columns 3 and 4, we examine *Exploit* and *Explore* indices to gauge innovation style. The results turn quite different as shown that the relative importance between inventor and firm fixed effects is about 1.4 in column 3 and about 1.3 in column 4, which suggest that the explanatory power of inventor fixed effects and firm fixed effects is very close. This result shows that both inventors and firms' time invariant characteristics are important determinants of innovation styles. Table 3 also reports the F-statistics for the joint significance of both fixed effects and the significance of fixed effects for inventor or firm individually. They all consistently fail to reject the null.

The explanatory power of all control variables and year effects are different when using different dependent variables: 28.3% for patent counts, 11.8% for patent citations, 7.4% for the *Exploit* index and 11.3% for the *Explore* index. Including inventor (or firm) fixed effects increases the adjusted R-squared. For example, adding firm fixed effects increases the adjusted R-square by 0.5% while adding inventor fixed effects increase the adjusted R-square by 6.6% when dependent variable is *LnPatent*. The extent of adjusted R-square increment corresponding to inventor (or firm) fixed effects is consistent with our results on the relative importance of inventor and firm fixed effects above. Coming back to the example of *LnPatent*, the ratio of the increment of adjusted R-square when including inventor fixed effects to the increment of adjusted R-square when including firm fixed effects is about 13, which is similar to our estimates

above.

3.3 Robustness checks

In this section, we conduct tests to check the robustness of our baseline findings.

First, we implement the MDV method used in Bertrand and Schoar (2003) by restricting the sample to the mobility sample in which only inventors who move across firms are included. The mobility sample includes 32,561 movers (21,139 movers when using *Exploit* and *Explore* indices as dependent variables) as well as the 4,310 (3,249 firms when using *Exploit* and *Explore* indices as dependent variables) firms for which they work. Note that in the mobility sample, the number of firms is equal to that in the connectedness sample because only firms with movers can be identified whatever the method adopted.

[Insert Table 4 Here]

Panel A of Table 4 reports the results using the MDV method. The inventor fixed effects continue to explain a significant portion of innovation performance. Specifically, inventor fixed effects have almost 7 to 8 times more explanatory power than firm fixed effects when using innovation performance measures (patent counts and citations) as dependent variables, which is consistent with the big gap of the explanatory power between inventor and firm fixed effects in our baseline regressions. In terms of the metrics of innovation styles (*Exploit* and *Exlpore 60*), the contribution of inventor fixed effects is similar to that of firm fixed effects in explaining the total variation as the ratio turns to be about 0.7, which is also consistent with our main results. Although the relative contribution between inventor and firm fixed effects is different when using the AKM method and the MDV method (e.g., changing from 13.0 to 7.7 when *LnPatent* is used, from 1.4 to 0.7 when *Exploit* is used), the main economic messages are the same with either the MDV method or the AKM method: inventors are way more important than firms when explaining innovation performance and inventors play a similar important role to firms when explaining innovation style. In fact, as detailed in Graham, Li, and Qiu (2012), such change in the relative contribution between inventor and firm fixed effects with two methods is the result of sample difference and normalization procedure. The incremental change in adjusted-R square across three different model specifications with the mobility sample are also consistent with our earlier findings reported in Table 3.

Second, we conduct another robust test using the AKM method on the largest group of

the connectedness sample. A noted issue in the AKM method is to pin down each connected group in the connectedness sample and all fixed effects are identified relative to a benchmark within each group. To ensure all fixed effects are comparable across groups in the connectedness sample, we follow Cornelissen (2008) to normalize both inventor and firm fixed effects. Nevertheless, there remains an issue regarding the change of the relative explanatory power of inventor and firm fixed effects as the normalization procedure re-weights the between-group explanatory power. To mitigate this concern, we re-estimate all our regressions using only the largest group of the connectedness sample and thereby normalization is not required. Moreover, the largest group is composed of 32,450 movers, 168,566 stayers, 4,130 firms, and 1,237,555 inventor-year observations, which accounts for 99% of inventors, 96% of firms, and 99% of inventor-year observations of the connectedness sample. This enlarged sample grants a great power to remedy the bias caused by normalization.

Panel B of Table 4 reports the results using the largest group and we find they are qualitatively similar to what we obtained in the baseline regressions. We reexamine the relative importance of inventor and firm fixed effects in determining inventors' innovation performance and innovation style. The AKM results using *LnPatent* and *LnCitation* as dependent variables from the largest group show that the ratio of inventor fixed effects' contribution to firm fixed effects' contribution in explaining total variation is about 13 to 14 times. When using *Exploit* and *Explore* as dependent variables to measure innovation style, the relative contribution between inventors and firms is about 1.5 times. All F-statistics for the joint significance of both fixed effects and the significance of fixed effects for inventor or firm individually consistently reject the null. We also report the adjusted R-square of the AKM method in all four columns. Overall, these additional tests suggest that our results in the baseline regressions are robust.

3.4 Heterogeneity of inventor fixed effects

So far, we have shown that inventor-specific effects explain a significant fraction of the variation in innovation performance and styles. Additionally, we would like to assess how big these observed differences inventors are. Therefore, we look at the distributions of fixed effects estimated above. We plot the density distribution function of estimated inventor fixed effects using four different dependent variables. Because fixed effects are estimated relative to a benchmark, the location and the mean of the estimated fixed effects may vary when picking

different benchmark. So we demean inventor fixed effects in all our figures without changing the shape of the distribution function.

[Insert Figure 1 Here]

In Figure 1, we plot four distributions in Panels A to D that corresponds to the retrieved inventor fixed effects with different dependent variables. Panels A and B of Figure 1 plot the estimated inventor fixed effects distributions when using metrics of innovation performance (*LnPatent* and *LnCitation*) as the dependent variables. Both distributions are slightly left skewed and this observation is consistent with the fact that patent count and citation data are left skewed due to the fact that many inventors filed no patents in some years. Our results contrasts with many prior studies that estimated individual fixed effects are roughly normally distributed (e.g. Graham, Li and Qiu (2012) and Ewens and Rhodes-Kropf (2015)), which underscores the features of invention: innovations, especially high-quality innovations, are mainly driven by a few inventors with great talent or inherent characteristics. It's also noteworthy that the distribution of estimated inventor fixed effects in terms of patent citation (Panel B) has higher dispersion and fatter tails than that of estimated inventor fixed effects in terms of patent number (Panel A). This observation indicates that the time invariant characteristics of inventors in determining the quality of innovations are much more dispersed than that of inventors in shaping the number of innovations. This estimate contributes to a large literature on innovation skill dispersion and productivities and could help parameterize models that proxy for variations in innovation skill (e.g., Iranzo, Schivardi and Tosetti (2006) and Bombardini, Gallipoli and Pupato (2012)).

In Panels C and D of Figure 1, we plot the distributions of estimated inventor fixed effects using the metrics of innovation style (*Exploit* and *Explore*) as dependent variables. Compared with Panels A and B, both distributions in C and D are more concentrated, which shows that there is a smaller difference in inventors' time invariant characteristics that determines their innovation style than that determines their innovation performance. Figure 2 shows the distribution of estimated inventor fixed effects in the largest group, which has a similar shape with Figure 1.

[Insert Figure 2 Here]

For a more precise assessment of the dispersion of inventor fixed effects, we tabulate the distribution of retrieved inventor fixed effects from the AKM regressions in Table 5. We show

median, standard deviation, 25th percentile, and 75th percentile in both Panel A and Panel B, in which Panel A corresponds to the distribution of inventor fixed effects using the connectedness sample and Panel B corresponds to the distribution of inventor fixed effects using the largest group.

[Insert Table 5 Here]

Overall, Table 5 shows that the variation in the size of inventor fixed effects is economically large. To highlight some examples, row 1 of Panel A shows that the difference between an inventor at the 25th percentile in the distribution of fixed effects in terms of patent counts and one at the 75th percentile is 0.45. In terms of patent citations in row 2 of Panel A, an inventor in the bottom quartile reduces the number of patent citation by 55% while an inventor in the top quartile increases the number of citation by 41%. For *Exploit* in row 3 of Panel A, the difference between the 25th percentile and 75th percentile is 0.12. Given that the average *Exploit* index in our sample is 0.13, the economic effect is large. Finally, in the last row of Panel A, the difference of fixed effects between an inventor in the 25th percentile and an inventor in the 75th percentile when using *Explore* is also large, amounting to 0.26.

4. Heterogeneity in the cross section

In this section, we conduct cross-sectional analysis in the full fixed effects model to better understand how inventor and firm heterogeneity alters our baseline results. These cross-sectional tests provide additional insights on how relative roles of inventors and firms would change in response to different characteristics of inventors and firms.

4.1 Firms with high centrality

In this subsection we examine whether the inventor is more important in firms with more "key" inventors. We define whether an inventor is a key inventor based on his or her normalized degree of centrality in the whole inventor co-authorship network. Following Hochberg, Ljungqvist, and Lu (2007), we calculate an inventor's normalized degree of centrality each year, which is equal to the number of coauthor relationships in the past three years an inventor has in the network divided by the maximum possible coauthor relationships in the past three years he or she can have in an n-inventor network. For example, in a co-authorship network with 10 inventors observed in the past three years, inventor $i's$ normalized degree of centrality is equal to

$n/9$ in which 9 is maximum number of ties inventor i can form in this network if he or she have coauthored with other inventors in the past three years. After computing all inventors' normalized degree of centrality in each year, we define an inventor of a year who is in the top 10% of normalized degree of centrality as a "key" inventor of the firm in that year. Then across our sample period, we calculate the average number of "key" inventors per year for each firm and select those connected firms that are in the top 10% in terms of the average number of "key" inventors. Our final subsample of firms with high centrality consists of 424 firms and 166,897 inventors when using *LnPatent* and *LnCitation* as the dependent variables, and 320 firms and 146,083 inventors when using *Exploit* and *Explore* as the dependent variables.

[Insert Table 6 Here]

We report the results of the AKM estimation in the subsample of firms with high centrality in Table 6. In each column, we report the contribution of inventor and firm fixed effects in explaining the variation in the subsample as well as the fraction of model R-square explained by each set of variables in parentheses. We also report the ratio of the contribution of inventor fixed effects to that of firm fixed effects.

Consistent with our main results, inventor fixed effects are more important than firm fixed effects in explaining innovation outcomes measured by patent counts, citations, *Exploit* or *Explore* indices. They contribute much more in determining innovation performance than innovation style. Furthermore, in a comparison with our baseline results the ratio of the contribution of inventor fixed effects to the contribution of firm fixed effects increases. For example, the ratio increases from 13.0 in the baseline regression to 15.8 in this subsample when using the patent counts as the dependent variable. The ratio increases from 1.4 to 1.5 when using *Exploit* as the dependent variable. In particular, the extent of the ratio increase is larger among regressions using the metrics of innovation performance. This result indicates that in firms with a higher degree of centrality (having more "key" inventors), the relative importance of inventors fixed effects in explaining innovation, in particular innovation performance, becomes more pronounced.

4.2 Firms with high mobility

We consider a subsample of firms with high mobility in this subsection to investigate the relation between inventor turnover of a firm and innovation output produced by inventors in this

firm. As discussed in the introduction, there are two competing arguments regarding the relative contribution of inventors in a firm with high mobility. On the one hand, high inventor turnover could imply that these firms' inventors are replaceable and hence firm organizational capital should matter more. One the other hand, high inventor turnover could indicate high standards implemented by the firm, which makes the firm pursue talented inventors and thereby induces larger inventor contribution to the firm's innovation output.

To examine these alternative arguments above, we construct a subsample of firms with high inventor mobility. We define the mobility of a firm as the ratio of total number of movers to the total number of inventors of the firm. We then pick those connected firms that are in the top 20% bracket in terms of their mobility scores to form our subsample.[8] The subsample is comprised of 556 firms and 20,691 inventors when using patent counts and citations as the dependent variable, and 433 firms and 17,611 inventors when using exploitative and explorative ratios as the dependent variable.

[Insert Table 7 Here]

Table 7 reports the results using the AKM method in the subsample of firms with high inventor mobility. In each column, we report the portions of inventor fixed effects and firm fixed effects that explains the variation in the subsample as well as the fraction of model R-square explained by each set of variables in parentheses. We also report the ratio of the contribution of inventor fixed effects to that of firm fixed effects.

The results suggest that inventor fixed effects are more important than firm fixed effects in explaining the four innovation metrics and inventor fixed effects contribute much more in determining innovation performance metrics than innovation style metrics. On top of that, different from our baseline results, the ratio of inventor fixed effect contribution to firm fixed effect contribution becomes larger in this subsample analysis. In column 1, the ratio rises from 12.0 (in the baseline tests) to 18.6 in the regression with patent counts as the dependent variable. In column 2 in which patent citation is the dependent variable, this ratio climbs rapidly from 14.3 (in the baseline tests) to 87.0. In the regressions with innovation style metrics are the dependent variable, this ratio increases from 1.4 to 2.0 in column 3 and from 1.3 to 2.3 in column 4. All these results show that, in firms with high inventor mobility (i.e., high frequency in inventor

---

[8] Here we do not restrict our sample to the top 10% as before because all these top 20% firms receive the same scores regarding mobility.

turnovers), the relative contribution of inventor is more important in explaining innovation outcomes, which supports our second arguments.

4.3 High-tech firms

The third dimension we consider when exploring heterogeneity focuses on a subsample of high-tech firms. Because human capital tends to be more important in high-tech industries, we postulate that inventor human capital plays a more significant role in determining innovation outcomes than firm organizational capital.

Following Tian and Wang (2014), we classify patents in our sample into four categories: (1) drugs, medical instrumentation, and chemicals (hereafter drugs); (2) computers, communications, and electrical (hereafter computers/electrical); (3) software programming and internet applications (hereafter software); (4) other miscellaneous patents. Based on the category of patents a firm produces most, we sort all our sample firms into one of above four categories. If a firm has no patent, then we classify it into one of these four categories based on the type of patents that are most frequently produced by the firm's industry. We then label all firms belonging to drugs and computers/electrical as high-tech firms and firms in software and miscellaneous as low-tech firms. Our subsample of high-tech firms consists of all firms in drugs and computers/electrical and inventors who work for these firms. The sample includes 2,556 firms and 145,857 inventors in the regressions using patent counts and citations as the dependent variable, and 1,996 firms and 132,157 inventors in the regressions using *Exploit* and *Explore* as the dependent variable.

[Insert Table 8 Here]

Table 8 presents the results using the AKM method to estimate both inventor fixed effects and firm fixed effects. In each column, we report the contributions of inventor and firm fixed effects in explaining the variation in the subsample and the fraction of model R-square explained by each set of variables in parentheses. We also compute the ratio of inventor and firm fixed effects' relative contribution.

Compared with our baseline results, these ratios in the high-tech subsample are larger: in column 1 the relative importance ratio increases from 13.0 to 15.4; in column 2 this ratio rises from 14.3 to 15.4; in column 3 the ratio increases from 1.4 to 1.6 and column 4 reports that the ratio increases from 1.3 to 1.6. These results suggest that in high-tech industries inventors'

human capital is more crucial in driving innovation outcomes than that of inventors in an average industry. Table 8 also reports that the results of F-tests for joint significance of both inventor fixed effects and firm fixed effects as well as the respective significance of each fixed effect. We are able to reject the null hypothesis.

4.4 Public Vs. private firms

Our final cross-sectional test compares public firms with private firms. Our baseline estimation and previous subsample tests focus on the connectedness sample containing only public firms because we match the inventor-year data with the financial data in Compustat. In this subsection, we include those unmatched private firms into our sample and re-estimate inventor fixed effects and firm fixed effects using the AKM method in the pooled sample including both public and private firms. We also separately do the same estimation in the private firm sample. Our goal is to explore the differences in the relative contribution of inventor human capital and firm organizational capital in explaining innovation outcomes between public and private firms.

Our pooled sample (including both public and private firms) has 63,501 firms and 431,903 inventors and the sample of private firms includes 51,454 firms and 280,001 inventors. It suggests that an average public firm has more inventors than private firms. Because we do not observe accounting information for private firms, our regressions exclude all firm financial controls. We, however, still control for inventors' prior innovation experience and year effects.

[Insert Table 9 Here]

Table 9 estimates the full fixed effects model using both the pooled sample of public and private firms and the subsample of private firms only. [9] In each column, we report the contributions of inventor fixed effects and firm fixed effects in explaining the variation in the sample and the fraction of model R-square explained by each set of variables in parentheses. We also compute the ratio of inventor and firm fixed effects' relative contribution.

We continue to observe that, in the pooled sample that includes both public and private firms, inventor fixed effects play a more significant role than firm fixed effects in driving

---

[9] As the sample size grows disproportionately after including private firms into our sample, the Stata command "felsdvreg" proposed by Cornelissen (2008) becomes infeasible in our analysis. Here we switch to another command "reghdfe" detailed by Correia (2014) to implement the AKM estimation, which is more efficient when dealing with data requiring large memory.

innovation. Further, inventors contribute more in determining innovation performance metrics (patent counts and citations) than innovation style metrics (exploitative and explorative ratios). For example, the ratio that assesses the relative contribution of inventor fixed effects and firm fixed effects is 9.3 in the pooled sample with patent counts as the dependent variable and it is 1.3 in the pooled sample with the exploitative score as dependent variable.

Next, we examine the difference between public and private firms in the relative contribution of inventor human capital and firm organizational capital. One argument is that public firms tend to attract more homogeneously talented inventors because they are more transparent and subject to a larger degree of pressures from short-term investors while private firms are more likely to hire more heterogeneous inventors due to more degree of freedom protected by their private ownership. This difference leads to the hypothesis that, in public firms, inventor human capital contributes more relative to firm organizational capital in determining innovation outcomes.

In Table 9, we compare the ratio of inventor fixed effects to firm fixed effects estimated from the pooled sample and from the private firm sample. In column 1 in which the dependent variable is patent counts, the ratio in the pooled sample, 9.3, is larger than that in the private sample, 7.5. In column 2 in which patent citation is the dependent variable, the ratio is 6.6 in the pooled sample and it is 5.3 in the private sample. Regarding innovation style metrics, we continue to observe that the ratio in the pooled sample is bigger than that in the private sample. These results suggest that inventors of private firms are relatively less important in determining innovation performance than that of public firms. The evidence is consistent with our conjecture.

5. Conclusion

In this paper, we have studied the relative importance of firm organizational capital and inventor human capital in promoting a firm's innovation output. We decompose the variation in innovation output into firm- and inventor-specific components. Inventor human capital appears to be about 13 times more important than firm organizational capital in explaining a firm's innovation performance in terms of patent counts and citations, while inventor human capital is only about two times more important when explaining the firm's innovation style in terms of patent exploratory and exploitive scores. In the cross section, inventors contribute more to innovation output when they are better networked, are in firms with higher inventor mobility, are

in high-tech industry firms, and are in publicly traded firms. This paper highlights the importance of individual inventors in enhancing firm innovation and sheds new light on the theory of the firm.

We need to bear in mind three caveats when interpreting or generalizing our results. First, similar to all other studies that use movements of individuals (e.g., executives, venture capitalists, bankers, employees, etc.) as an identification strategy, our empirical setting is still subject to the concern that moving of inventors could be endogeneous. Our results intend to show the average effect of moving across firms on inventors who actually move. We are silent about the reason for inventors' moving. Second, we are only able to capture the contribution to innovation output from movers and stayers in firms with at least one mover. If there is no inventor moving in a firm, we would not be able to separate the contribution from the inventors and that from the firms. Finally, because innovation is human capital intensive, we are likely to attribute more innovation contribution to inventors. Hence, what we find is likely a lower bound of a firm's contribution to investment.

References

Abowd, J. M., R. H. C., and F. Kramarz, 2002. Computing person and firm effects using linked longitudinal employer-employee data, Technical Report 2002-06. *U.S. Census Bureau*.

Abowd, J. M., F. Kramarz, P. Lengermann, and S. Perez-Duarte, 2004. Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States. *Unpublished working paper*.

Abowd, J. M., F. Kramarz, and D. N. Margolis, 1999. High wage workers and high wage firms. *Econometrica* 67(2), 251-333.

Acharya, V. and K. Subramanian, 2009. Bankruptcy codes and innovation. *Review of Financial Studies* 22, 4949-4988.

Acharya, V. V., R. P. Baghai, and K. V. Subramanian., 2014. Wrongful discharge laws and innovation. *Review of Financial Studies* 27, 301-346.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt, 2005. Competition and innovation: An inverted-U relationship. *Quarterly Journal of Economics* 120, 701-728.

Aghion, P., J. Van Reenen, and L. Zingales, 2013. Innovation and institutional ownership. *American Economic Review* 103, 277-304.

Andrews, M. J., L. Gill, T. Schank, and R. Upward, 2008. High-wage workers and low-wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171, 673-697.

Bertrand, M., and A. Schoar, 2003. Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics* 118(4), 1169-1208.

Berk, J. B., J. H. Binsbergen, and B. Liu, 2014. Matching capital and labor. *Unpublished working paper*.

Bombardini, M., G. Gallipoli, and G. Pupato, 2012. How robust is the skill-dispersion-complementarity hypothesis? *American Economic Review* 102(5), 2327-2348.

Brav, A, J. R. Graham, C. R. Harvey, and R. Michaely, 2005. Payout policy in the 21st century. *Journal of Financial Economics* 77(3), 483-527.

Chemmanur, T. J., M. Ertugrul, and K. Krishnan, 2015. Is it the investment bank or the investment banker? A study of the role of investment banker human capital in acquisitions. *Unpublished working paper*.

Coase, Ronald H., 1988, The nature of the firm: Origin, meaning, influence. *Journal of Law, Economics, and Organization* 4, 3-17.

Cornaggia, J., Y. Mao, X. Tian, and B. Wolfe, 2015. Does banking competition affect innovation? *Journal of Financial Economics* 115(1), 189-209.

Cornelissen, T., 2008. The Stata command felsdvreg to fit a linear model with two high-dimensional fixed effects. *Stata Journal* 8(2), 170.

Correia, S., 2014. REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. *Statistical Software Components*.

Ewens, M., and M. Rhodes-Kropf, 2015. Is a VC partnership greater than the sum of its partners? *Journal of Finance* 70(3), 1081-1113.

Gao, H., P. Hsu, and K. Li, 2015. Public market listing and corporate innovation strategies: Evidence from private firms. *Unpublished working paper*.

Graham, J. R., S. Li, and J. Qiu, 2012.Managerial attributes and executive compensation. *Review of Financial Studies* 25, 144-186.

Griliches, Z., 1990. Patent statistics as economic indicators: a survey. *Unpublished working paper*.

Griliches, Z., A. Pakes, and B. Hall, 1988. The value of patents as indicators of inventive activity. *Unpublished working paper.*

Hall, B. H., A. B. Jaffe, and M. Trajtenberg, 2001. The NBER patent citation data file: Lessons, insights and methodological tools. *Unpublished working paper*.

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

Hart, O., 1995. Firms, contracts, and financial structure. *Oxford University Press*.

He, J. J., and X. Tian, 2013. The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics* 109(3), 856-878.

Hochberg, Y. V., A. Ljungqvist, and Y. Lu, 2007. Whom you know matters: Venture capital networks and investment performance. *Journal of Finance* 62(1), 251-301.

Hsu, Po-Hsuan, X. Tian, and X. Yan, 2014. Financial development and innovation: Cross-country evidence. *Journal of Financial Economics* 112, 116–135.

Iranzo, S., F. Schivardi, and E. Tosetti, 2008. Skill dispersion and firm productivity: An analysis with employer-employee matched data. *Journal of Labor Economics* 26(2), 247-285.

Kaplan, S. and L. Zingales, 1997. Do financial constraints explain why investment is correlated with cash flow? *Quarterly Journal of Economics* 112, 169-215.

Katila, R., and G. Ahuja, 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal* 45, 1183-1194.

Klein, B., 1988. Vertical integration as organizational ownership: The fisher body-general motors relationship revisited. *Journal of Law, Economics and Organizations* 4, 199-213.

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

Manso, G., 2011. Motivating innovation. *Journal of Finance* 66, 1823-1860.

Seru, A., 2014. Firm boundaries matter: Evidence from conglomerates and R&D activity. *Journal of Financial Economics* 111(2), 381-405.

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

Sørensen, J. B., and T. E. Stuart, 2000. Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly* 45(1), 81-112.

Steven N. K., B. A. Sensoy, and P. Stromberg, 2009. Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies. *Journal of Finance* 64(1), 75-115.

Tian, X., and T. Y. Wang, 2014. Tolerance for failure and corporate innovation. *Review of Financial Studies* 27(1), 211-255.

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

Williamson, O. E., and S. G. Winter, eds., 1993. The nature of the firm, origins, evolutions, and development. *Oxford University Press*.

Wooldrige, J. M., 2001. Econometric analysis of cross-section and panel data. *MIT Press*.

Zingales, L., 2000. In search of new foundations. *Journal of Finance* 55, 1623–1653.

Appendix A

Details in sample construction

 We match the HBS patent and inventor database with the NBER patent citation database following four steps:

 (1) We break all patents in the NBER database into two subsets based on the number of assignees each patent has---one subset (hereafter called Subset A) contains all patents owned by a single assignee while the other one (called Subset B) includes patents owned by multiple assignees. For inventors whose patents belong to Subset A, their company affiliations can be easily identified. We match all patents in Subset A with the HBS database through patent number, resulting in a set of 6,270,074 matched inventor-patent instances denoted by Counterpart A.

 (2) We divide the set, Counterpart B, consisting of all unmatched instances after step one into two groups: one group referred to as Counterpart B1 that is comprised of inventor-patent instances in which each patent is filed by a single inventor; and the other group referred to as Counterpart B2 that collects the rest inventor-patent instances in which every patent is coauthored by multiple inventors. We then match all instances in Counterpart B1 with Subset B by patent number, which leads to 22,555 inventor-patent-assignee instances corresponding to 11,461 inventor-patent instances in Counterpart B1 as each patent may be possessed by several assignees. We then determine one assignee for each instance based on matched information in Counterpart A, i.e., we designate a unique assignee to an inventor in the year that patent granted if this assignee coincides with one assignee for which the inventor has been recognized to work in Counterpart A in the same year. If we dig out multiple assignees through above method, we exploit the location information to pin down the assignee for these instances. Another extreme case is that we find no appropriate assignees in Counterpart A through above method, we also exploit the location information to help us determine the assignee for these instances. For example, if the inventor's location is Mahwah of New Jersey, the assignee with perfectly matched location would be assigned. If several assignees have the same location, we randomly choose one. Otherwise, we relax our searching criteria and select the assignee sharing the same state, New Jersey and so on so forth. In such way, we can figure out all assignees for the 11,461 instances.

 (3) For all instances in Counterpart B2, they are patents filed by multiple inventors and belonged to various assignees. Through patent number, we join them with Subset B to form all

pairwise combinations and then select one assignee for each inventor-patent instance. The selection procedure is identical to that in step 2. As a result, assignees for the 250,168 inventor-patent instances in Counterpart B2 can be identified.

(4) Combing all instances obtained in above three steps, our final matched sample consists of 6,531,703 inventor-patent instances whose assignee can be uniquely identified.

Then we nail down the company affiliation for each inventor over his/her whole career with the assistance of 6,531,703 matched inventor-patent instances. If all patents filed by an inventor of a year belong to a single assignee, it can be ascertained that the inventor was hired by the assignee in that year with certainty. Another situation often encountered is that patents filed by an inventor of one year are owned by different assignees. For instance, two patents of an inventor are claimed by assignee A while another five patents belong to assignee B in a year. In such case, it's quite reasonable to assume that the inventor was employed by the assignee to which most patents of the inventor belong in that year. Particularly, when the number of patents to different assignees breaks even, say an inventor files the same number of patents for both assignee A and B in certain year, we utilize the inventor's employment information of last year to help us identify---if he worked for assignee A (B) last year, we presume that he was employed by A (B) this year in order to make his career path consistent.[10] Otherwise, we randomly pick one for him. This procedure brings about 4,251,546 inventor-year observations.

For our analysis, we augment our inventor-year sample in a chronological order by filling all year gaps for inventors who appear in the patent database but do not have patents in the gap years. For example, an inventor filed patents in 1986 and 1991. Thus our sample only captures the inventor's performances in 1986 and 1991. We expand the observations between 1986 and 1991 for him or her by taking the patent counts and patent citations to be zero.[11] This method comes with a caveat: that how we can accurately identify inventors' employer in gap years. Following above example, it would be quite intuitive and easy for us to decide which company the inventor belongs to between 1986 and 1991 if the patents he filed in both years are owned by the same company. It would turn to be difficult if the patents filed in 1986 and 1991 belong to

---

[10] Admittedly, this is an ad hoc assignment. To alleviate this concern, we repeat our analysis after different assigning methods. For example, we use the inventor's employment information in subsequent year, i.e., if the inventor worked for assignee A (B) in next year, we assume his employer of this year was A (B). We also tried the method to pick an assignee for the inventor randomly. Doing these does not alter the nature of the results.

[11] However, we specify other metrics that evaluate inventors' innovation style (such as exploratory ratio and exploitative ratio elaborated in next subsection) as missing.

distinct companies, say, A and B. In other words, how do we decide the company affiliations of a mover for the transition years where we have no observations on his patent filing? We adopt the method by assuming that the inventor belongs to the old company A in the first half of his transition years (1987 to 1988) and belongs to the new company B in the second half of the transition years (1989 to 1990).[12] This procedure leads to 7,445,855 inventor-year observations in our augmented sample.

---

[12] We tried other methods to assign the company affiliation to a mover: (1) assume the inventor belongs to the old company A for all the transition years; (2) assume the inventor belongs to the new company B for all the transition years. Also, we conducted our analysis in which we excluded all years with no observations on patent filing. The results are qualitatively similar across these implementations.

Appendix B

Definition of variables

Variable	Definition
Measures of innovation	
$LnPatent_{t+3}$	Natural logarithm of one plus inventor i's total number of patents filed (and eventually granted) in year $t+3$;
$LnCitePat_{t+3}$	Natural logarithm of one plus inventor i's total number of citations received on the inventor's patents filed (and eventually granted), scaled by the number of the patents filed (and eventually granted) in year $t+3$;
$Exploit_{t+3}$	The number of exploitative patents filed (and eventually granted) divided by the number of all patents filed (and eventually granted) by the inventor in year $t+3$; a patent is classified as exploitative if at least 60% of its citations are based on existing knowledge;
$Explore_{t+3}$	The number of exploratory patents filed (and eventually granted) divided by the number of all patents filed (and eventually granted) by the firm in year $t+3$; a patent is classified as exploratory if at least 60% of its citations are based on new knowledge;
Measures of control variables	
$LnExpnum_{t+3}$	Natural logarithm of one plus inventor i's average number of patents filed (and eventually granted) per year in the three years prior to year $t+3$;
$LnExpcit_{t+3}$	Natural logarithm of one plus inventor i's average number of scaled citations received on the inventor's patents filed (and eventually granted) in three years prior to year $t+3$;
$Assets_t$	Book value of total assets (#6) measured at the end of fiscal year t;
$R\&DAssets_t$	Research and development expenditures (#46) divided by book value of total assets (#6) measured at the end of fiscal year t, set to 0 if missing;
Age_t	Firm i's age, approximated by the number of years the firm has been listed on Compustat;
ROA_t	Return-on-assets ratio defined as operating income before depreciation (#13) divided by book value of total assets (#6), measured at the end of fiscal year t;
$PPEAssets_t$	Property, Plant & Equipment (net, #8) divided by book value of total assets (#6) measured at the end of fiscal year t;
$Leverage_t$	Firm i's leverage ratio, defined as book value of debt (#9 + #34) divided by book value of total assets (#6) measured at the end of fiscal year t;
$CapexAssets_t$	Capital expenditure (#128) scaled by book value of total assets (#6) measured at the end of fiscal year t;
$TobinQ_t$	Firm i's market-to-book ratio during fiscal year t, calculated as [market value of equity (#199 $\times$ #25) plus book value of assets (#6) minus book

	value of equity (#60) minus balance sheet deferred taxes (#74, set to 0 if missing)] divided by book value of assets (#6);
KZindex_t	Firm *i*'s KZ index measured at the end of fiscal year *t*, calculated as -1.002 $\times$ Cash Flow ((#18+#14)/#8) plus 0.283 $\times$ Q ((#6+#199$\times$#25-#60-#74)/#6) plus 3.189 $\times$ Leverage ((#9+#34)/(#9+#34+#216)) minus 39.368 $\times$ Dividends ((#21+#19)/#8) minus 1.315 $\times$ Cash holdings(#1/#8), where #8 is lagged;
Hindex_t	Herfindahl index of 4-digit SIC industry *j* where firm *i* belongs, measured at the end of fiscal year *t*.

Panel A

Distribution of Estimated Inventor Fixed Effects: LnPatent

Panel B

Distribution of Estimated Inventor Fixed Effects: LnCitation

Panel C

Distribution of Estimated Inventor Fixed Effects: Exploit60

Panel D

Distribution of Estimated Inventor Fixed Effects: Explore60
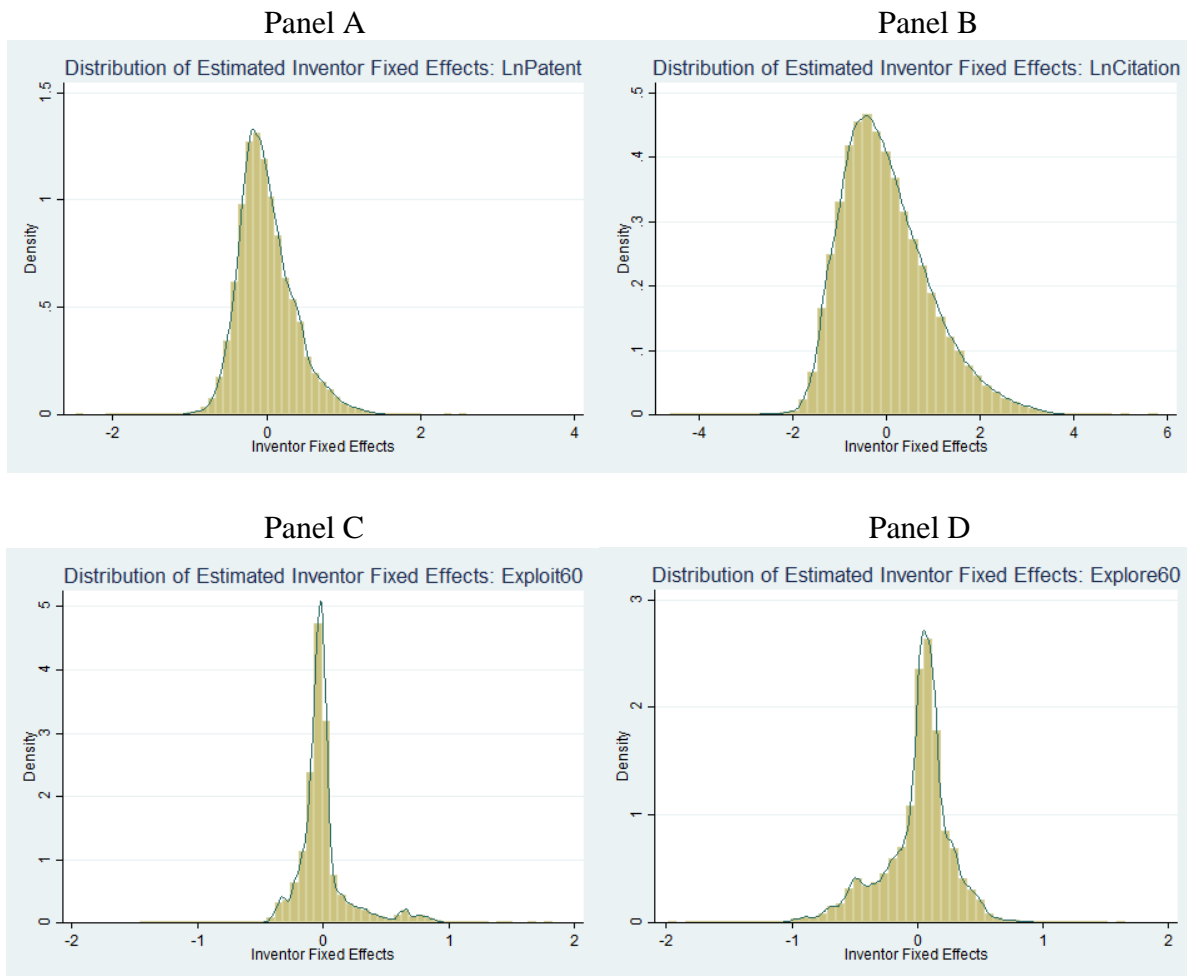
Figure 1. Distribution of estimated inventor fixed effects: connectedness sample. This figure displays the distribution of estimated inventor fixed effects from the AKM regression in the connectedness sample with four different dependent variables: the log of one plus the adjusted number of patents (Panel A), the log of one plus the adjusted number of patent citations (Panel B), the *Exploit* index (Panel C), and *Explore* (Panel D). In the graph, the estimates are normalized so the mean value of the inventor fixed effects is zero.
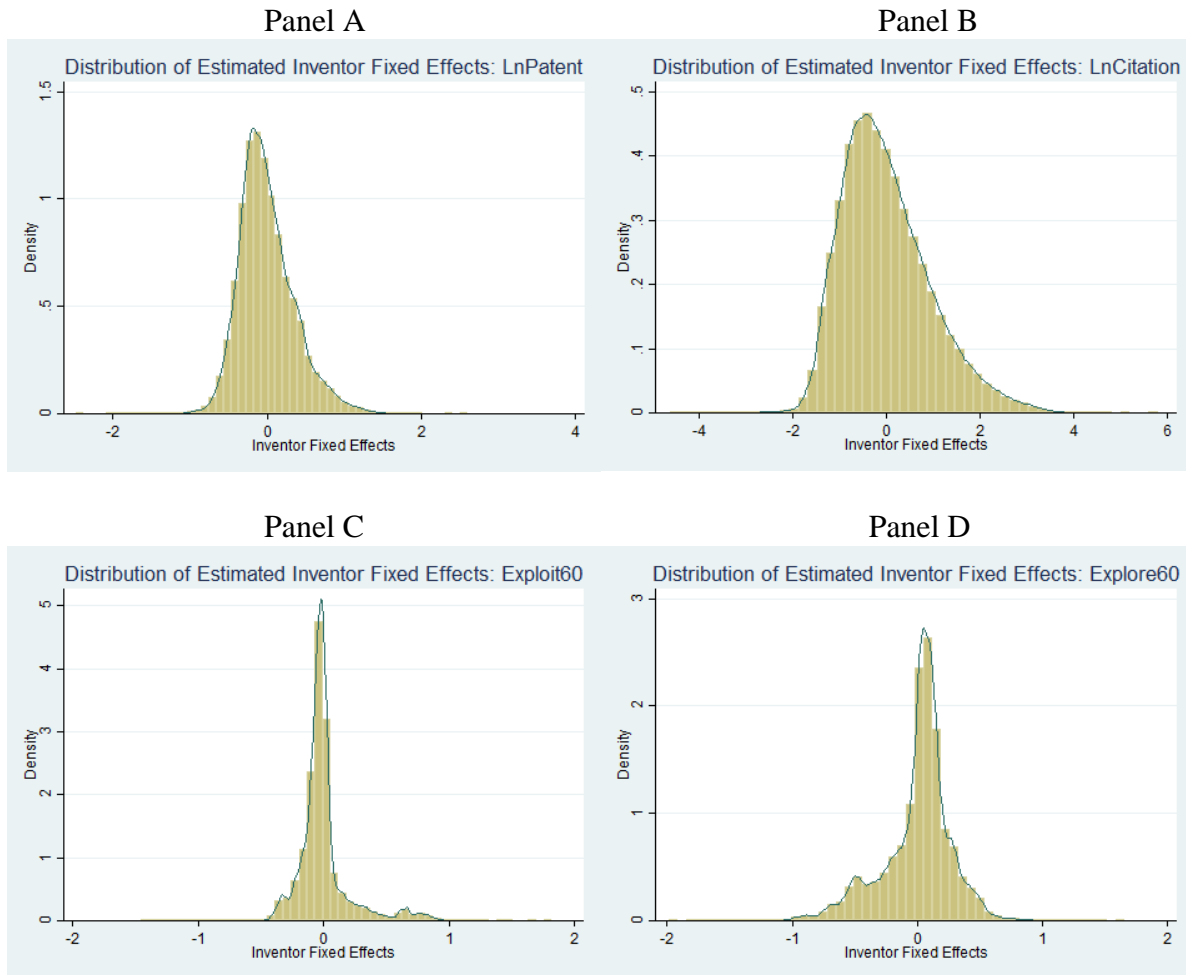
Figure 2. Distribution of estimated inventor fixed effects: largest group of the connected sample. This figure displays the distribution of estimated inventor fixed effects from the AKM regression in the largest connected sample with four different dependent variables: the log of one plus the adjusted number of patents (Panel A), the log of one plus the adjusted number of patent citations (Panel B), the Exploit index (Panel C), and Explore (Panel D). In the graph, the estimates are normalized so the mean value of the inventor fixed effects is zero.

Table 1: Inventor mobility characteristics—full sample

This table reports the mobility structure of our full sample from 1970 to 2006. A mover is defined as the inventor who switches the firm for which he or she works for. Panel A presents the employment information of inventors in our sample. Panel B presents the distribution of movers across firms.

Panel A: Number of mover out of all inventors

Mover	No. of firms in which inventors are employed	No. of inventors	Percent
No	1	172,117	84.09
	Subtotal	172,117	84.09
	2	28,159	13.76
	3	3,776	1.84
Yes	4	552	0.27
	5	65	0.03
	6	7	0.00
	7	2	0.00
	Subtotal	32,561	15.91
	Total	204,678	100.00

Panel B: Number of movers per firm

Mover Per Firm	Frequency	Percent	Cum.
0	1,412	24.68	24.68
1-5	1,719	30.04	54.72
6-10	734	12.83	67.55
11-20	610	10.66	78.21
21-30	274	4.79	83.00
31-50	268	4.68	87.68
51-100	253	4.42	92.10
>100	452	7.90	100.00
Total	5,722	100.00	

Table 2: Summary statistics

This table reports summary statistics for the full sample and the connectedness sample in inventor and firm level when using patent counts and citations as dependent variables in our baseline regression. Panel A presents the summary statistics of patent counts and citations as well as the time-varying inventor characteristics measures for inventor-year observations. Panel B presents the summary statistics of the time-varying firm characteristics measures for firm-year observations. Definitions of variables are listed in the table of our appendix.

Panel A: Summary statistics for inventors

Variable	Mean	Median	SD	25th	75th	N (inventor-year)
Patent						
Full Sample	0.965	0.000	2.100	0.000	1.000	1,250,041
Connectedness Sample	0.967	0.000	2.105	0.000	1.000	1,239,614
Citation						
Full Sample	6.899	0.000	17.020	0.000	7.456	1,250,041
Connectedness Sample	6.914	0.000	17.044	0.000	7.472	1,239,614
Expnum						
Full Sample	0.850	0.333	1.643	0.000	1.000	1,250,041
Connectedness Sample	0.852	0.333	1.648	0.000	1.000	1,239,614
Expcit						
Full Sample	7.186	2.927	12.713	0.000	9.121	1,250,041
Connectedness Sample	7.199	2.939	12.717	0.000	9.146	1,239,614

Panel B: Summary statistics for firms

Variable	Mean	Median	SD	25th	75th	N (firm-year)
Assets (million)						
Full Sample	7120.888	691.710	21416.620	110.870	4084.982	46,177
Connectedness Sample	7569.452	844.164	22253.390	137.228	4512.000	40,047
RDAssets						
Full Sample	0.059	0.032	0.143	0.010	0.068	46,177
Connectedness Sample	0.060	0.035	0.140	0.012	0.069	40,047
Age						
Full Sample	21.932	21.000	13.151	10.000	32.000	46,177
Connectedness Sample	22.386	22.000	13.305	11.000	33.000	40,047
ROA						
Full Sample	0.115	0.143	0.268	0.090	0.195	46,177
Connectedness Sample	0.121	0.145	0.254	0.093	0.197	40,047
PPEAssets						
Full Sample	0.298	0.269	0.161	0.189	0.377	46,177
Connectedness Sample	0.299	0.271	0.158	0.193	0.377	40,047
Leverage						
Full Sample	0.222	0.209	0.176	0.101	0.309	46,177
Connectedness Sample	0.220	0.208	0.173	0.103	0.304	40,047
CapexAssets						
Full Sample	0.068	0.059	0.047	0.038	0.085	46,177
Connectedness Sample	0.069	0.060	0.046	0.039	0.086	40,047
TobinQ						
Full Sample	1.940	1.317	2.635	1.021	1.967	46,177
Connectedness Sample	1.942	1.339	2.463	1.034	2.005	40,047
KZindex						
Full Sample	-5.429	-1.656	88.358	-4.482	-0.065	46,177
Connectedness Sample	-4.997	-1.714	52.882	-4.518	-0.129	40,047
Hindex						
Full Sample	0.263	0.213	0.187	0.128	0.345	46,177
Connectedness Sample	0.264	0.214	0.187	0.127	0.348	40,047

Table 3: Inventor and firm fixed effects—Connectedness sample regressions

This table reports three-way fixed effects regressions using the method in Abowd, Kramarz and Margolis (1999) and Abowd, Creecy, and Kramarz (2002) to estimate both inventor and firm fixed effects in the connectedness sample. The estimation is implemented by exploiting the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the log of one plus the adjusted number of patents as the dependent variable and column 2 uses the log of one plus the adjusted number of patent citations as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 exploit the *Exploit* and *Explore* indices as the dependent variables (missing value is assigned if no patents filed by an inventor of a year). The rows for "Relative Importance of Inventor and Firm Fixed Effects" report the inventor and firm fixed effects in explaining the total variation in the connectedness sample. The percentages in parentheses present the fraction of model R-square explained by each set of variables (see Graham, Li, and Qiu (2012), Ewens and Rhodes-Kropf (2015) for details). The row for "Inv. FE / Firm FE" calculates the ratio of the contribution of inventor fixed effects to the contribution of firm fixed effects in explaining the total variation. The rows for "F-test on Fixed Effects" report the F-statistics for the joint significance of both fixed effects and respective significance of inventor and firm fixed effects. For each column, the rows for "Adj. R-sq." report three regressions with different set of independent variables. The first row reports adjusted R^2 for the regression including all time-varying control variables. The second row reports adjusted R^2 for the regression including all control variables and firm fixed effects. The third row reports adjusted R^2 for the regression including all control variables and inventor fixed effects. The last row reports adjusted R^2 for the AKM estimation. Definitions of variables are defined in Appendix B. "# Firms" is the total firms in our sample. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	LnPatent	LnCitation	Exploit	Explore
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.311 (67.76%)	0.300 (90.36%)	0.340 (56.95%)	0.326 (51.34%)
Contribution of Firm FE	0.024 (5.45%)	0.021 (6.33%)	0.236 (39.53%)	0.245 (38.58%)
Inv. FE / Firm FE	12.958	14.286	1.441	1.331
F-test on Fixed Effects				
Joint F-statistic	1.63***	1.61***	2.54***	2.79***
Inventor FE F-statistic	1.59***	1.54***	1.31***	1.37***
Firm FE F-statistic	1.65***	1.59***	3.83***	5.22***
Observation	1,239,614	1,239,614	548,233	548,233
Adj. R-sq. after the Addition of Inventor and Firm Fixed Effects				
Control Variables	0.283	0.118	0.074	0.113
Adding Firm FE	0.288	0.128	0.328	0.379
Adding Inventor FE	0.349	0.197	0.376	0.427
Adding Both FE	0.351	0.199	0.391	0.448
# Movers	32,561	32,561	21,139	21,139
# Stayers	168,900	168,900	161,940	161,940
# Firms	4,310	4,310	3,249	3,249
Year Effects	Yes	Yes	Yes	Yes

Table 4: Robust checks

This table reports robust check results using the MDV method detailed in Bertrand and Schoar (2003) to estimate both inventor and firm fixed effects in the mobility sample (Panel A) and results using the AKM method to estimate both inventor and firm fixed effects in the largest group of connected sample (Panel B). In both panels, column 1 uses the log of one plus the adjusted number of patents as the dependent variable and column 2 uses the log of one plus the adjusted number of patent citations as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 exploit the *Exploit* and *Explore* indices as the dependent variables (missing value is assigned if no patents filed by an inventor of a year). Definitions of variables are defined in the table of appendix. In Panel A, "# Person" reports the number of movers since only movers are included in the mobility sample. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

Panel A: Inventor and firm fixed effects based on the mobility sample

	(1)	(2)	(3)	(4)
Dependent Variable	LnPatent	LnCitation	Exploit	Explore
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.254 (59.62%)	0.226 (79.58%)	0.190 (41.21%)	0.179 (35.52%)
Contribution of Firm FE	0.033 (7.75%)	0.032 (11.27%)	0.257 (55.75%)	0.247 (49.01%)
Inv. FE / Firm FE	7.697	7.063	0.739	0.725
F-test on Fixed Effects				
Joint F-statistic	1.89***	1.73***	2.80***	2.94***
Inventor FE F-statistic	1.89***	1.67***	1.24***	1.30***
Firm FE F-statistic	1.45***	1.43***	4.99***	5.49***
Observation	323,381	323,381	113,072	113,072
Adj. R-sq. after the Addition of Inventor and Firm Fixed Effects				
Control Variables	0.287	0.125	0.048	0.104
Adding Firm FE	0.293	0.137	0.281	0.332
Adding Inventor FE	0.348	0.187	0.219	0.271
Adding Both FE	0.352	0.192	0.313	0.368
# Persons (#Movers)	32,561	32,561	21,139	21,139
# Firms	4,310	4,310	3,249	3,249
Year Effects	Yes	Yes	Yes	Yes

Panel B: Inventor and firm fixed effects based on the largest group

Dependent Variable	(1) LnPatent	(2) LnCitation	(3) Exploit	(4) Explore
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.310 (67.69%)	0.300 (90.36%)	0.340 (56.86%)	0.325 (51.26%)
Contribution of Firm FE	0.025 (5.46%)	0.021 (6.33%)	0.237 (39.63%)	0.245 (38.64%)
Inv. FE / Firm FE	12.400	14.286	1.435	1.327
F-test on Fixed Effects				
Joint F-statistic	1.63***	1.61***	2.54***	2.79***
Inventor FE F-statistic	1.59***	1.54***	1.31***	1.37***
Firm FE F-statistic	1.68***	1.61***	3.93***	5.28***
Observation	1,237,555	1,237,555	546,829	546,829
Adj. R-sq.	0.351	0.199	0.391	0.448
# Movers	32,450	32,450	20,997	20,997
# Stayers	168,566	168,566	161,392	161,392
# Firms	4,130	4,130	3,055	3,055
Year Effects	Yes	Yes	Yes	Yes

Table 5: Distribution of retrieved inventor fixed effects

This table tabulates distributions of retrieved inventor fixed effects from the AKM regressions using four different dependent variables in both the connectedness sample (Panel A) and the largest group of connected sample (Panel B). The estimates are normalized so the mean value of the inventor fixed effects is zero.

Panel A: Summary statistics of inventor fixed effects in the connectedness sample

Dep. Variable	Median	SD	25th	75th	Number of Inventor
LnPatent	-0.052	0.361	-0.245	0.200	201,461
LnCitation	-0.138	0.947	-0.687	0.556	201,461
Exploit	-0.029	0.211	-0.090	0.025	183,079
Explore	0.047	0.279	-0.108	0.147	183,079

Panel B: Summary statistics of inventor fixed effects in the largest group

Dep. Variable	Median	SD	25th	75th	Number of Inventor
LnPatent	-0.052	0.361	-0.245	0.200	201,016
LnCitation	-0.138	0.947	-0.687	0.556	201,016
Exploit	-0.029	0.210	-0.090	0.025	182,389
Explore	0.047	0.279	-0.108	0.147	182,389

Table 6: Subsample of firms with high centrality

This table reports the subsample analysis results using the AKM method to estimate both inventor and firm fixed effects in a subsample of firms with high centrality. We define the centrality of firm based on their employers' degree centrality. From 1970 to 2003, we calculate inventors' normalized degree centrality of a year, which is equal to the number of coauthor relationships of past three years an inventor has in the network divided by the maximum possible coauthor relationships of past three years he or she could have in an n-inventor network (Refer to Hochberg, Ljungqvist, and Lu (2007) for more detail on the methodology). In each year, we define an inventor who is in the top 10% of normalized degree centrality as a "key" inventor of the firm. Then across our sample period, we compute the average number of "key" inventors per year for all firms and select those connected firms that are in the top 10% in terms of the average number of "key" inventors, which gives us a subsample of firms with high centrality. The estimation is implemented by exploiting the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the log of one plus the adjusted number of patents as the dependent variable and column 2 uses the log of one plus the adjusted number of patent citations as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 exploit the *Exploit* and *Explore* indices as the dependent variables (missing value is assigned if no patents filed by an inventor of a year). Definitions of variables are defined in the table of appendix. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	LnPatent	LnCitation	Exploit	Explore
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.315 (68.18%)	0.303 (91.54%)	0.346 (58.05%)	0.327 (51.58%)
Contribution of Firm FE	0.020 (4.33%)	0.014 (4.23%)	0.230 (38.59%)	0.245 (38.64%)
Inv. FE / Firm FE	15.750	21.643	1.504	1.335
F-test on Fixed Effects				
Joint F-statistic	1.70***	1.60***	2.67***	2.96***
Inventor FE F-statistic	1.66***	1.54***	1.39***	1.43***
Firm FE F-statistic	4.86***	3.24***	17.86***	21.48***
Observation	1,032,675	1,032,675	452,755	452,755
Adj. R-sq.	0.358	0.202	0.402	0.459
# Movers	17,688	17,688	11,037	11,037
# Stayers	149,209	149,209	135,046	135,046
# Firms	424	424	320	320
Year Effects	Yes	Yes	Yes	Yes

Table 7: Subsample of firms with high mobility

This table reports the subsample analysis results using the AKM method to estimate both inventor and firm fixed effects in a subsample of firms with high mobility. The subsample of firms with high mobility includes only the set of connected firms in top 20% (actually these firms receive same scores regarding mobility) of mobility which is equal to the ratio of the total number of movers a firm has to the total number of inventors of that firm. The estimation is implemented by the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the log of one plus the adjusted number of patents as the dependent variable and column 2 uses the log of one plus the adjusted number of patent citations as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 exploit the *Exploit* and *Explore* indices as the dependent variables (missing value is assigned if no patents filed by an inventor of a year). Definitions of variables are defined in the table of appendix. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	LnPatent	LnCitation	Exploit	Explore
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.334 (72.29%)	0.348 (98.31%)	0.370 (62.93%)	0.365 (57.48%)
Contribution of Firm FE	0.018 (3.90%)	0.004 (1.13%)	0.182 (30.95%)	0.156 (24.57%)
Inv. FE / Firm FE	18.556	87.000	2.033	2.340
F-test on Fixed Effects				
Joint F-statistic	1.47***	1.44***	1.63***	1.68***
Inventor FE F-statistic	1.44***	1.37***	0.94	1.12***
Firm FE F-statistic	1.51***	1.77***	5.73***	5.97***
Observation	99,885	99,885	42,289	42,289
Adj.R-sq.	0.316	0.175	0.282	0.363
# Movers	3,699	3,699	2,715	2,715
# Stayers	16,992	16,992	14,896	14,896
# Firms	556	556	433	433
Year Effects	Yes	Yes	Yes	Yes

Table 8: Subsample of high-tech firms
   This table reports the subsample analysis results using the AKM method to estimate both inventor and firm fixed effects in a subsample of high-tech firms. The subsample of high-tech firms considers only the set of firms in "Drug & Chemical" category that includes industries mainly producing patents on drugs, medical instrumentation, and chemicals or in "Computer & Electrical" category that includes industries mainly producing patents on computers, communications technologies, and electrical technologies (See Tian and Wang (2014) for details). The estimation is implemented by the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the log of one plus the adjusted number of patents as the dependent variable and column 2 uses the log of one plus the adjusted number of patent citations as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 exploit the *Exploit* and *Explore* indices as the dependent variables (missing value is assigned if no patents filed by an inventor of a year). Definitions of variables are defined in the table of appendix. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	LnPatent	LnCitation	Exploit	Explore
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.323 (70.07%)	0.308 (92.22%)	0.360 (59.90%)	0.343 (54.53%)
Contribution of Firm FE	0.021 (4.56%)	0.020 (5.99%)	0.224 (37.27%)	0.210 (33.39%)
Inv. FE / Firm FE	15.381	15.400	1.607	1.633
F-test on Fixed Effects				
Joint F-statistic	1.59***	1.58***	2.70***	2.79***
Inventor FE F-statistic	1.56***	1.52***	1.29***	1.37***
Firm FE F-statistic	1.76***	1.65***	4.00***	4.62***
Observation	874,399	874,399	403,168	403,168
Adj.R-sq.	0.351	0.199	0.401	0.444
# Movers	19,815	19,815	13,555	13,555
# Stayers	126,042	126,042	118,602	118,602
# Firms	2,556	2,556	1,996	1,996
Year Effects	Yes	Yes	Yes	Yes

Table 9: Public firms Vs. private firms

This table reports the subsample analysis results using the AKM method to estimate both inventor and firm fixed effects in a sample including both public and private firms beginning from 1970 to 2003. Private firms are those firms unmatched with Compustat Annual Data while public firms are identical to those appeared in our baseline regressions. The estimation is implemented by the Stata command "reghdfe" proposed by Correia (2014). In our regressions, all time-varying firm characteristics are dropped as private firms have no financial data, whereas time-varying inventor characteristics (*Prior Innovation Experience*) are included. Each column corresponds to two sets of regression using different samples: one regresses on the pooled sample while the other one regresses on only the sample of private firms. Column 1 uses the log of one plus the adjusted number of patents as the dependent variable and column 2 uses the log of one plus the adjusted number of patent citations as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 exploit the *Exploit* and *Explore* indices as the dependent variables (missing value is assigned if no patents filed by an inventor of a year). Definitions of variables are defined in the table of appendix. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

	(1)		(2)		(3)		(4)	
Dependent Variable	LnPatent		LnCitation		Exploit		Explore	
Sample	Pooled	Private	Pooled	Private	Pooled	Private	Pooled	Private
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)								
Contribution of Inv. FE	0.260	0.276	0.250	0.267	0.260	0.280	0.260	0.292
	(62.65%)	(66.03%)	(84.18%)	(83.44%)	(54.28%)	(61.54%)	(50.88%)	(58.99%)
Contribution of Firm FE	0.028	0.037	0.038	0.050	0.195	0.147	0.205	0.161
	(6.75%)	(8.85%)	(12.79%)	(15.63%)	(40.71%)	(32.31%)	(40.12%)	(32.53%)
Inv. FE / Firm FE	9.286	7.459	6.579	5.340	1.333	1.905	1.268	1.814
F-test on Fixed Effects								
Joint F-statistic	1.65***	1.60***	1.67***	1.63***	2.64***	1.98***	2.87***	2.29***
Observation	3,624,251	2,071,875	3,624,251	2,071,875	1,302,755	661,081	1,302,755	661,081
Adj.R-sq.	0.323	0.307	0.187	0.184	0.307	0.239	0.349	0.295
# Inventors	431,903	280,001	431,903	280,001	293,754	165,926	293,754	165,926
# Firms	63,501	51,454	63,501	51,454	29,420	213,29	29,420	213,29
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes