Cornell University School of Hotel Administration

The Scholarly Commons

Working Papers School of Hotel Administration Collection

5-2017

The Role of Human Capital: Evidence From Patent Generation

Tong Liu
University of Pennsylvania

Yifei Mao
Cornell University School of Hotel Administration, ym355@cornell.edu

Xuan Tian
Tsinghua University

The Role of Human Capital: Evidence From Patent Generation

Abstract

Firms exhibit persistence in innovation output. This paper focuses on the role played by individual inventors. Compared to firm organizational capital, human capital embedded in inventors explains a majority of the variation in innovation performance but much less in innovation style. Inventors contribute more when they are better networked, in firms with higher inventor mobility, and in industries in which innovation is more difficult. Additional tests suggest that our main findings are unlikely driven by inventors' endogenous moving. This paper highlights the importance of human capital in enhancing firm innovation and sheds new light on the theory of the firm.

Keywords

human capital, organizational capital, inventors, innovation persistence, JEL G30, JEL G32, JEL O32

Disciplines

Finance and Financial Management

Comments

The role of human capital: Evidence from patent generation

Tong Liu
University of Pennsylvania
tongl@wharton.upenn.edu
(215) 898-7622

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

Xuan Tian
Tsinghua University
tianx@pbcsf.tsinghua.edu.cn
(86) 10-6279-4103

This version: May, 2017

*We are grateful for the comments and suggestions from Linda Canina, Daniel Ferreira, Andrew Ellul, Gerard Hoberg, Andrew Karolyi, Pam Moulton, Veronica Pool, Noah Stoffman, Scott Yonker, Alminas Zaldokas, and seminar and conference participants at Cornell University, Indiana University, the 2016 SFS Cavalcade conference, the 2016 European Finance Association annual meetings, and the 2016 China International Conference in Finance. We alone are responsible for any errors or omissions.

The role of human capital: Evidence from patent generation

Abstract

Firms exhibit persistence in innovation output. This paper focuses on the role played by individual inventors. Compared to firm organizational capital, human capital embedded in inventors explains a majority of the variation in innovation performance but much less in innovation style. Inventors contribute more when they are better networked, in firms with higher inventor mobility, and in industries in which innovation is more difficult. Additional tests suggest that our main findings are unlikely driven by inventors' endogenous moving. This paper highlights the importance of human capital in enhancing firm innovation and sheds new light on the theory of the firm.

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

JEL number: G30, G32, O32

1. Introduction

Since Schumpeter, existing literature has well established that innovation is crucial to firms' competitive advantages (Porter, 1992) and long-run growth of the economy (Solow, 1957). The differential ability to generate innovation, argued by prior studies, is the major source of heterogeneity in firm performance (Nelson, 1991, 1995; Winter, 1984, 2003). However, there are large differences in innovation output between top innovative and the rest of the firms. For example, Cohen et al. (2013) report persistence and predictability in a firm's ability to do R&D and produce patents.

What drives a firm's persistence in innovation output? There must be something time-invariant and unique embedded in the firm. Since Coase (1937), there has been a longstanding debate on the unique factor that defines a firm. In the Hart-Moore framework, nonhuman assets, such as physical assets, organizational structure, corporate culture, and access to resources, are the glue that brings a firm together (see, e.g., Hart (1995) for a detailed discussion). Zingales (2000), however, argues that "human capital is emerging as the most crucial asset" in today's world and highlights the importance of human capital. It is an open yet unanswered question on the relative importance of firm organizational capital and human capital of the employees inside the firm in terms of contributing to the firm's innovation persistence. In this paper, we attempt to answer this research question and shed new light on the understanding of what constitutes a firm.

A major challenge of our study is to separate the contribution of organizational capital and human capital. Patent generation (by individual inventors) provides a unique setting that is clean and rich when dealing with this empirical challenge. First, innovation typically requires specialized knowledge and skills possessed by researchers (instead of top executives). Therefore, inventors play a crucial role in driving a firm's innovation output. By focusing on inventors, we capture the key human capital in a firm. Second, we are able to track individual inventors' patent filings and the corresponding patent assignees. As a result, we are able to observe inventors' move from one firm to another, using information on patent assignees.[1] These two unique features of our setting allow us to separate the contribution of firm organization capital and inventor human capital. Intuitively, if an individual inventor's output does not change as she moves, the inventor is likely to be the main driver of her output. If, however, an individual

[1] Patents are typically assigned to the firm by which the inventor is employed.

inventor's innovation output changes significantly after she moves from one firm to another, then we can largely attribute the change in her innovation output to the new firm she is joining.

We begin our analysis by formally documenting innovation persistence in a framework that is different from Cohen et al. (2013). Taking advantage of patent inventor data, we find evidence on innovation persistence at both inventor and firm levels. Specifically, inventors with higher innovation output in the past tend to possess higher patent counts or citations in the future. For example, a one standard deviation increase in inventor patent filings in the current year is associated with a 12.8% increase in patent filings in the next year, and a one standard deviation increase in citations per patent in the current year is associated with 12.5% increase in citations per patent next year. Given the skewness in patent filings, the strength of persistence in innovation is sizable. Consistent with Cohen et al. (2013), we find firm-level innovation persistence as well.

The above findings highlight the importance of inventors (as well as firms) in generating innovation output, and raise an important yet unanswered question – the relative importance of inventors and firms. To isolate the roles played by inventors and firms, we use two methods. First, we focus on a panel of inventors who have changed their affiliated firms, and include inventor-, firm-, and year-fixed effects in the regressions. We refer to this approach as the mover dummy variable (henceforth MDV) method, which has been commonly used in the existing literature (e.g., Bertrand and Schoar, 2003; Graham, Li, and Qiu, 2012). Because the MDV approach is limited to movers (i.e., inventors who change their affiliated firms in our sample period) only, which accounts for 16% of all inventors in our sample, we use a second method. This approach includes both movers and stayers (i.e., inventors whose affiliated firms remain unchanged) in the sample, as long as the stayers are in the firms that employ 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 a rather small sample of movers to a connectedness sample, which includes 98.4% of all inventors in our setting. Both methods allow us to quantify how much of the observed heterogeneity in patent generation can be attributed to inventor fixed effects and firm fixed effects. The main results and their economic implications are similar regardless of the method we use.

We investigate two dimensions of innovation output in this paper. The first dimension is innovation performance, measured by both the quantity and quality of a firm's innovation output.

The second dimension is innovation "style", measured by the exploratory score and exploitative score of a firm's innovation output.[2] We find that inventor fixed effects are consistently more important than firm fixed effects in explaining innovation performance. However, they are about the same important in explaining innovation style. Specifically, our estimates suggest that inventor fixed effects are about 6 times as important as firm fixed effects in explaining a firm's innovation performance, while inventor fixed effects are about the same as important as firms fixed effects when explaining the firm's innovation style. The results suggest that while inventor human capital is crucial in determining a firm's innovation performance, its role on a firm's innovation style is much more moderated.

Furthermore, we find that the relative importance of inventor human capital to firm organizational capital exhibits significant heterogeneity in the cross section. First, we examine inventor network. The degree of an inventor's centrality is determined by the number of coauthor relationships she has. Compared to an average inventor in terms of network centrality, better connected inventors contribute significantly more to innovation performance but not innovation style than firms do.

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 driving forces. On the one hand, high inventor turnover may imply that these inventors are replaceable and hence firms' organizational capital matters more. One the other hand, high inventor turnover could indicate that the firm implements high standards and pursues talented inventors, which induces larger inventor contribution to the firm's innovation output. We find that the second effect dominates: Inventors in high mobility firms appear more important in determining innovation performance (but not style) compared to inventors in low mobility firms.

Third, we explore industry heterogeneity and focus on industries in which innovation is more difficult to achieve. It is likely that in these industries inventor human capital plays a more dominanting role than firm organizational capital. We find that, in drug, chemical, computer, and electrical industries that are typically considered as high-tech industries, inventors are more crucial in driving innovation performance (but not style) than those in other industries.

[2] As defined in the existing literature (e.g., Levinthal and March, 1993; McGrath, 2001; Benner and Tushman, 2002; Smith and Tushman, 2005; Gao, et al., 2015), exploratory innovation is radical innovation that requires knowledge outside of the existing knowledge domain, and exploitative innovation is incremental innovation that builds on existing knowledge and improves existing skills, processes, and structures.

In the final part of the paper, we conduct a series of tests to address the concern that inventors' endogenous moving across firms could bias our estimation. As the entire existing literature uses the moving of individuals (e.g., CEOs, venture capitalists, bankers, fund managers) to isolate the effects of human capital from organizational capital and there is not a method that can perfectly address the endogenous moving issue, we gauge the extent to which potential biases caused by endogenous moving could be. We first show that the firms to which inventors move are on average smaller, younger, and have worse operating performance than the firms from which inventors leave. We also find that movers are on average less productive than stayers. These results suggest that inventors' moving is more likely to be involuntary (and hence exogenous) and their moving is likely to be caused by their own poor innovation productivity in the firms they leave. Hence, endogenous moving by inventors does not appear a major concern.

Next, to the extent that endogenous moving still remains a concern, we gauge how much it would bias our main results. We first compare the relative importance of inventors and firms in a subsample that contains inventors who "move up" to a firm with better operating performance (hence the moving is more likely to be voluntary and endogenous) and in a subsample that contains inventors who "move down" to a firm with worse operating performance (hence the moving is more likely to be involuntary and exogenous). We find that the relative importance of inventors tofirms in explaining innovation performance is higher for moving-up inventors. However, the relative importance in explaining innovation style is similar across the two groups of inventors. The results suggest that, while there is some overestimation of firms' contribution for moving-up inventors, endogenous moving by inventors does not appear to substantially alter our main results.

To address the concern that firm organizational capital might be underestimated if inventors endogenously choose to move to similar firms in which they are less likely to experience changes in innovation output, we compare the relative importance of inventors and firms when inventors move to similar firms (i.e., with similar past operating performance or in the same industry) and when inventors move to firms that are very different from their current firms (i.e., firms with different past operating performance or in a different industry). We find that the relative importance of inventors to firms in explaining innovation performance and style is indeed higher in the firms whose inventors move to similar firms. Nonetheless, our results from inventors who move to different types of firms once again suggest that our main findings are not entirely driven by inventors' endogenous moving decisions.

Finally, we examine the relative importance of inventors and firms in a subsample that is less likely to be subject to the matching concern. Endogenous matching between inventors and firms could cause us to wrongly attribute the change in the inventor's output to firm organizational capital, where the change actually comes from an interaction between the inventor and the firm. In other words, matching might contaminate our estimation of the relative importance of the inventor and the firm. To address this concern, we follow Graham, Li, and Qiu (2012) and focus on a subsample of movers who exhibit little changes in their innovation output as they move across firms. The rationale of this test is that these movers are less likely to be subject to endogenous matching issues. Our main findings continue to hold in this test.

While we have performed a rich set of tests to assess the degree of potential biases caused by endogenous moving to the best we can, one caveat is that we cannot completely rule out the matching concern due to inventors' endogenous moving. This issue is unfortunately a common challenge that the existing strand of labor and finance literature faces. While our main findings – human capital plays a more important role in explaining innovation performance and a relatively less important role in explaining innovation style – survive a rich set of robustness tests, one still needs to be cautious when interpreting or generalizing our results.

The 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 capital from human capital (Coase, 1988; Klein, 1988; Williamson and Winter, 1993). An empirical study by Kaplan et al. (2009) examines startup companies and suggests that business (nonhuman capital) is more important than management team (human capital). Different from their work, our paper focuses on established firms and highlight the importance of human capital in patent generation in these firms. One plausible reason for our different results is that we explore established firms that are very different in nature and many other aspects from startup companies. Hence, our paper supplements Kaplan et al. (2009).

Second, our paper contributes to the innovation literature by documenting the importance of inventors' human capital in explaining innovation performance and style. To the best of our knowledge, our paper is the first one that attempts to isolate the contribution of inventor human capital from firm organizational capital in determining a firm's innovation output. Prior studies have examined various 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 only explore on the extensive margin regarding the determinants of firm-level innovation output. In this paper, we delve further into the intensive margin and decompose innovation drivers into human-capital- and organization-capital-related components, which allow us to further understand the relative importance of these two components in a firm's patent generation.

Third, this paper 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 (2012) show that manager fixed effects explain a significant extent of firm policy heterogeneity, and managers with higher performance fixed effects receive higher compensation. Ewens and Rhodes-Kropf (2015) 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. Cho et al. (2016) find that firms are more important than managers in driving patent generation.[3] Existing studies, however, are either unable to capture individuals' output in the setting of workers or firm/fund managers, or have to infer individuals' output through indirect ways (such 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 analyze individual inventors' performance and style.

The rest of the paper is organized as follows. Section 2 describes data and variable constructions. Section 3 documents firm innovation persistence. Section 4 reports the main results. Section 5 presents subsample results that explore the heterogeneity of our main findings. Section 6 addresses concerns on inventors' endogenous moving. Section 7 concludes and discusses potential caveats of our results.

---

[3] Cho et al. (2016)'s findings are likely driven by the fact that mangers are not the main drivers of patent generation. While the executives may set the boundaries for the R&D group, the exact innovation directions and outputs are determined by the knowledge and expertise possessed by inventors in the firm.

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.[4] 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 an inventor as she moves from one firm to another. Since patent database does not directly indicate the employment of an 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 the 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.[5] We discuss in more details regarding how to pin down the employer for each inventor when there are multiple assignees in Appendix A.

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 are only a small portion of our final sample. Our final sample consists of 204,678 inventors (1,246,951 inventor-year observations) that have worked for 5,722 firms from 1970 to 2003.[6]

2.2 Variable measurement

2.2.1 Measuring Innovation

[4] See Li et al. (2014) and Singh and Fleming (2010) for more details about the HBS patent and inventor database.
[5] This database is available at https://sites.google.com/site/patentdataproject/Home/downloads. See Hall, Jaffe, and Trajtenberg (2001) for more details about the NBER patent citation database.
[6] The time span of our inventor sample ends in 2003 because we construct the innovation metrics three years ahead in our analysis to capture the long-term nature of innovation activity.

We construct two sets of patent-based metrics to gauge an inventor's innovation output. The first set measures innovation performance and the second set captures 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 her innovation output. 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 construct, 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 an inventor's innovation output.

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 which is updated till 2006). To adjust the truncation bias in patent counts, we calculate the number of patents filed by each inventor (and eventually granted) 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. However, patents tend to receive citations over a long period after its grant date, but we observe the citations received up to 2010. 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.

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 (*LnCitePat)* as the main innovation metrics to measure innovation performance in our analysis. We also winsorize all our dependent variables at the 99^{th} percentile.

The second set of metrics includes patents' exploitative (*Exploit*) and exploratory scores (*Explore*), which reflect an inventor's innovation style. We follow existing literature (e.g., Sorensen and Stuart, 2000; Katila and Ahuja, 2002; etc.) to categorize an inventor's patenting 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 her previous patent portfolio and the set of patents that has been cited by her own patents over the past five years. We then construct proxies so that a patent is classified as being exploitative if at least 60% of its citations are based on existing knowledge, and a patent is classified as being exploratory if at least 60% of its citations are based on new knowledge.[7] We then set *Exploit* equal to the ratio of the number of exploitative patents filed by an inventor in a given year to the total number of patents filed by the same 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, the patent 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* in 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 constraints (Kaplan and Zingales (1997)'s five-variable KZ index), and industry concentration (the Herfindahl index based on sales). Aghion et al. (2005) point out the non-linear effect 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.

[7] We use 80% as a cutoff, too. The results are robust to using variables that are defined by this alternative cutoff.

2.2.2 Summary statistics

Table 1 presents information on the movers and stayers in our sample. Panel A shows that during the sample period, 15.9% of (or 32,561) inventors are movers who work in more than one firm in the sample, while the rest 84.1% 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. 24.7% of the sample firms do not have any movers while the remaining 75.3% of 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 analysis on the mobility sample, which comprises of 32,561 movers as well as 4,310 firms at which these movers are employed.

2.2.3 Measuring control variables

Following the innovation literature, we control for a vector of inventor and firm characteristics that may affect innovation output. 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 innovation metrics (adjusted patent count and citations per patent, 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). The construction of these variables 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.

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, which examines whether the connectedness sample is representative of the full sample (Brav et al., 2005).[8] Panel A summarizes the representativeness of these variables

[8] 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.

for inventors. On average, an inventor in the sample has 0.9 granted patents per year and each patent receives 6.5 citations. In the connectedness sample, an inventor has a similar number of patents granted per year, 0.9, and each patent receives 6.5 citations. The other variables that measure how exploratory or exploitative an inventor could be are also very close in both the full sample and the connectedness sample.

Panel B of Table 2 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. PERFORMANCE PERSISTENCE OF INVENTORS AND FIRMS

Persistence has been documented on various dimensions such as capital structure (e.g. Denis, 2012; Leary and Roberts, 2005, etc.), operating performance (Denis and McKeon, 2016), and venture capital performance (Ewens and Rhodes-Kropf, 2015). Performance persistence implies that past performance can predict future outcomes in our setting. Thus, we concentrate on the relation between the inventor's or firm's past performance and their achievement of the current year. Our conjecture is that they would perform better at current year if they had better performance in the past. To gauge the performance of inventors and firms, we mainly focus on the number of patents filed and corresponding patent citations each year. We evaluate the inventor's and firm's performance history based on two measures: the cross-sectional ranking of their performance each year and patent counts or citations per patent each year.

Simple regressions would be subject to the concern that poor-performing inventors or firms are more likely to drop from the sample and hence what one observes in the data is just those who survive. In such cases, this approach would lead to misleading conclusions of positive performance persistence. To address this concern, in our regressions we only consider those inventors and firms with long span in our sample. As a result, it's unlikely that their performance persistence is attributed to pure luck that cannot last for a long time. In particular, we focus on the set of inventors who have at least 10-year observations and the set of firms that have at least 5-year observations in our sample.[9] Therefore, our performance persistence tests are based on inventors and firms that have appeared in our sample for a long time.

We construct two performance measures in this test. The first measure is relative performance ranking, we create a dummy variable that equals one if the inventor or firm is in the top quartile in terms of the number of filed patents (or patent citations) each year and zero otherwise. [10] The second measure is the direct count of patent numbers (or patent citations) filed by inventors or firms each year. A feature of the direct counting is that it's not stationary because in most years a lot of inventors in our sample have no patent filed. To overcome the nonstationarity problem, we use the past *t* years' average number of filed patent (or patent citations) to predict the current year's number of patents (or patent citations) they file and check whether persistence emerges in this specification.

3.1 Persistence in rankings

Table 3 reports the results of pooled regressions that regress inventor (Panels A and B) and firm (Panels C and D) relative rankings at the current year on those *t* year ago, *Top_t* (where $t = 1, 2, \text{ or } 3$). Panels A and C report the results with patent counts while Panels B and D report the results with patent citations. In inventor regressions, controls include both year and individual fixed effects. In firm regressions, we include all financial controls for firms described in Appendix A as well as year and firm fixed effects.

The coefficient estimates on *Top_t* are positive and significant at the 1 % level in all columns, suggesting a strong, positive relation between the inventor's (and the firm's) past and current performance ranking. For example, in Panel A, column (1) suggests that for an inventor who is in the top quartile one year ago, she is 13.1% more likely to stay in the top quartile in terms of patent counts in the current year, compared to inventors who are not in the top quartile one year ago. This result remains robust after we include *Top_2* or *Top_3* in the regressions. Similarly, according to our results in columns (2) and (3) of Panel A, an inventor who is in the top quantile 2 years (3 years) ago, they are 8.5% (2.9%) more likely to keep in the top quartile in

[9] The results are qualitatively similar if we use alternative cutoffs (e.g., 8 or 12 years for inventors, 4 or 6 years for firms) when selecting the samples.
[10] Our results are quite similar if we use top percentile, top quintile, or top tercile to define the ranking dummy.

terms of patent counts in the current year. In Panel B, when we use patent citations, our main results continue to hold. Panel C reports firm performance persistence results in terms of patent counts. The coefficient estimates on *Top_t* are all positive and significant. The economic magnitude is sizeable as well. For example, column (1) suggests that a firm that is in the top quartile in terms of patent counts one year again is 40.3% more likely to stay in the top quartile in the current. Panel D reports similar findings when we alternatively use a firm's ranking in terms of patent citations.

One potential concern is that inventors and firms operating in different industries might not be comparable in terms of patent counts and citations because patenting might be easier in some industries than that in others. To alleviate this concern, we define our top quartile inventors and firms within the industry in each year and repeat all our tests. The results are similar to those reported in Table 3.

3.2 Persistence in counts

In this subsection, we use alternative measures, patent counts and citations (as opposed to rankings), and explore whether the positive performance persistence remains. As we mentioned earlier, the number of filed patents is not a stationary series. To address this issue, we calculate *NPat_t* (where t = 1, 2, 3, or 4) as the average number of filed patents among previous t years by an inventor (firm).[11] Also, to avoid serial correlations of error terms in overlapping sample regressions, we construct a non-overlapping sample in the sense that we move our event windows accordingly without overlapping observation for inventors (firms).

Table 4 reports our regression results. The coefficient estimates on *NPat_t* are positive and significant at the 1% level in columns (1) to (4) of Panel A, which suggests a positive correlation between an inventor's past average number of filed patents and current patent counts. We obtain a similar result on inventor performance persistence in terms of patent citations, as Panel B shows. Regarding the results on firm performance persistence, Panels C and D report the results. The coefficient estimates on *NPat_t* are positive and significant in all columns, suggesting that a firm's past patent counts (and citations) could predict its current innovation output.

[11] At the firm level, the nonstationarity property of patent variables might not hold. For consistency, however, we run regressions by exploiting the identical method as we do with the inventor data.

4. Main results

Having formally documented innovation persistence at both the firm and inventor level, in this section, we explore the relative importance of firm organizational capital and inventor human capital that contributes to a firm's innovation output.

4.1 Empirical methods

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

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

where i denotes inventor, j denotes firm, and t denotes year when the innovation activity occurs. In the above equation, X_{it} and Z_{jt} include time-varying controls for inventor and firm. The time-varying controls at the firm level include firm asset size, R&D expenditure, age, profitability (measured by ROA), tangible assets (measured by PPE), leverage, capital expenditure, Tobin's Q, financial constraints, and product market competition. The time-varying controls at the inventor level include inventors' previous innovation performance (measured by patent counts and citations per patent in previous three years). μ_t captures the year fixed effects. γ_{it} controls for the functional area effects that inventor i belongs in year t. We define the functional area for an inventor as the primary functional class of most patents that inventor i applies for in year t.[12] Our focus is to retrieve both inventor and firm fixed effects ϕ_i and θ_j using 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 in a 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: We start

[12] The patent classes are based on the USPTO 3-digit classification system as of 2006. We assign the class of the largest number of patents that inventor *i* applies for in year *t* as his/her primary functional area. We supplement the functional area of inventor *i* with the class of the largest number of patents the firm (he/she works for) filed in year *t* if the inventor does not apply for any patent of that year. Similarly, for firms with no patents filed, we assign the functional area of the firm as its industry functional area (very few in our sample).

with an arbitrary inventor and track all firms where she has ever worked. Then we include all inventors whoever work in these firms into our connectedness sample and continue tracking all firms for which these inventors have ever worked. We repeat the procedure until all data are exhausted.[13] 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 the 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 a larger sample size and higher statistical power. Second, it can significantly reduce the computational work in terms of the large data set used in our study. Nonetheless, we 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 in the connectedness sample. Define the variable D_{ikt} as a dummy that equals one if inventor i works at firm k at time t and zero otherwise. Then we can rewrite equation (1) as:

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

[13] 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. 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.

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

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

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:

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

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 the above regression, we can recover the inventor fixed effects by the following equation:

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

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 because our sample contains about 16% movers. This proportion is pretty high compared to previous literature (e.g. Graham, Li, and Qiu, 2012). Another property of the AKM estimator is that fixed effect estimates themselves have properties that are similar to other estimators. As shown by Wooldridge (2010), the estimates of the time-varying variable coefficients are both unbiased and consistent, while the fixed effect estimates are only unbiased.

4.2 Baseline results

In this section, we analyze how unobserved inventor and firm time invariant characteristics affect inventors' innovation performance and style using the AKM method. The AKM method uses the connectedness sample that excludes firms that do not have a mover during our sample period. Based on this procedure, the connectedness sample 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.[14]

We follow the prior literature to select the observable characteristics of inventors and firms that may affect an inventor's future innovation output (e.g., He and Tian, 2013; 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 assets, and on inventor time-varying variables, such as prior experience of an inventor. Additionally, we include year fixed effects to capture the impact of economic conditions.

Table 5 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 capture the contribution of inventor fixed effects to the total variation in inventors' innovation output. $\frac{cov(Y,Inventor\ FE)}{var(Y)}$ 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 individual fixed effects in explaining the dependent variable for a given regression model. In addition, we report adjusted R-squared across four different model specifications in Table 5: the first specification includes all control variables and year effects; the second specification includes all control variables, year effects, and firm fixed effects; the third specification includes all control variables, year fixed effects, and inventor fixed effects; the final specification includes all control variables and all fixed effects. We adopt adjusted R-squares in this case because the number of explanatory variables changes across models.

For patent counts in column 1, inventor fixed effects account for 53.1% of the total variation while firm fixed effects contribute to 8.4% 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

[14] The final connected sample that enters into AKM estimation is 201,803 inventors and 4,294 firms. It is because functional area effects are missing for some of the inventors.

around 6 times in column 1. For citations per patent in column 2, 62.2% of the total variation corresponds to inventor fixed effects while 7.2% of the total variation corresponds to firm fixed effects. The relative importance between inventor and firm fixed effects amounts to about 8 times. Overall, the stark differences in explanatory power between inventor and firm fixed effects for patent counts and citations per patent 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 output.

In columns 3 and 4, we examine exploitative and exploratory scores to gauge innovation style. The results show 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 indicates that the explanatory power of inventor fixed effects and firm fixed effects is very close in explaining innovation style. The result suggests that the firm's organizational capital has a relatively more important impact on innovation style than innovation performance. While inventors appear to be able to carry their innovation ability to the new firm they move to, their innovation style is more likely to be affected by the new environment they get into. Table 5 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 reject the null that these coefficient estimates are jointly zero.

The explanatory power of all control variables except for inventor fixed effects and year fixed effects vary with different dependent variables: 35.4% for patent counts, 21.3% for citations per patent, 11.7% for the exploitative score and 15.6% for the exploratory score. Including inventor (or firm) fixed effects increases the adjusted R-squared. For example, adding firm fixed effects increases the adjusted R-squared by 1.9 percentage points while adding inventor fixed effects increases the adjusted R-squared by 8.4 percentage points when dependent variable is *LnPatent*. The extent of adjusted R-squared increment corresponding to inventor (or firm) fixed effects is consistent with our results on the relative importance of inventor and firm fixed effects above. For the example of *LnPatent*, the ratio of the increment of adjusted R-squared when including inventor fixed effects to the increment of adjusted R-squared when including firm fixed effects is about 4.4, which is close to our estimates above.

4.3. Robustness checks

In this subsection, we conduct additional tests to check the robustness of our baseline findings.[15]

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 32,420 movers (21,133 movers when using *Exploit* and *Explore* indices as dependent variables) as well as 4,294 (3,249 firms when using *Exploit* and *Explore* indices as dependent variables) firms for which they work. Note that the number of firms in the mobility sample equals that in the connectedness sample because only firms with movers can be identified no matter we use the AKM or MDV method.

Panel A of Table 6 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 4 times more explanatory power than firm fixed effects when using innovation performance measures (patent counts and citations per patent) as the dependent variable, 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 innovation style variables (*Exploit* and *Explore*), 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 6.3 to 4.7 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 but inventors play a similar important role with firms when explaining innovation style. In fact, as detailed in Graham, Li, and Qiu (2012), such a 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-squared across different model specifications with the mobility sample is also consistent with our earlier findings reported in Table 5.

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 that fixed effects are identified relative to a benchmark within each group. To ensure all fixed effects are comparable across

[15] More additional robustness tests and supplemental analyses are provided in Internet Appendix.

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 the largest group of the connectedness sample only 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 account for 99% of inventors, 96% of firms, and 99% of inventor-year observations of the connectedness sample, respectively. This sample grants a great power to mitigate the bias caused by normalization.

Panel B of Table 6 reports the results using the largest group, and the results 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 *LnCitePat* as the dependent variable from the largest group show that the ratio of inventor fixed effects' contribution to firm fixed effects' contribution in explaining total variation is about 6 times. When using *Exploit* and *Explore* as the dependent variable to measure innovation style, the relative contribution between inventors and firms is about 1.4 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-squared of the AKM method in all four columns. Overall, these additional tests suggest that our results in the baseline regressions are robust.

4.4 Heterogeneity in inventor fixed effects

So far, we have shown that inventor-specific effects explain a significant fraction of the variation in innovation performance and style. Additionally, we would like to assess how big these observed differences among inventors are. Therefore, we look at the distributions of inventor fixed effects. 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 using different benchmarks. So we demean inventor fixed effects in all our figures without changing the shape of the distribution function.

In Figure 1, we plot four distributions in Panels A to D that correspond 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 *LnCitePat*) as the dependent variables. Both distributions are slightly right skewed and this observation is consistent with the fact that patent count and citation data is right skewed due to the fact that many inventors file no patents in some years. Our results differ from many prior studies that find individual fixed effects to be roughly normally distributed (e.g. Graham, Li, and Qiu, 2012; Ewens and Rhodes-Kropf, 2015) because of a unique feature of invention: innovation, especially high-quality innovation, is 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 citations per patent (Panel B) has higher dispersion and fatter tails than those of estimated inventor fixed effects in terms of patent counts (Panel A). This observation indicates that time-invariant characteristics of inventors in determining the quality of innovations are much more dispersed than those of inventors in shaping the number of patents. This estimate contributes to a large literature on innovation skill dispersion and productivity, and could help parameterize models that proxy for variation in innovation skills (e.g., Iranzo, Schivardi, and Tosetti, 2006; Bombardini, Gallipoli, and Pupato, 2012).

In Figure 1 Panels C and D, we plot the distributions of estimated inventor fixed effects using the metrics of innovation style (*Exploit* and *Explore*) as the dependent variable. Compared with Panels A and B, both distributions in Panels C and D are more concentrated, which show that there is a smaller difference in inventors' time invariant characteristics that determine their innovation style than that determine 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.

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 7. We show median, standard deviation, 25th percentile, and 75th percentile in both Panel A and Panel B, in which Panel A shows the distribution of inventor fixed effects in the connectedness sample and Panel B shows the distribution of inventor fixed effects in the largest group.

Overall, Table 7 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 of the distribution on the natural logarithm of patent counts and the one at the 75th percentile of the distribution is 0.33. The difference is equivalent to 0.32 patents, which is 36% of our patent count sample mean (0.91). In terms of the natural logarithm of citations per patent, row 2 of Panel A shows that the difference between an inventor at the 25th percentile and one at the 75th percentile is 0.694. The difference is equivalent to 0.90 citations per patent, which accounts for 14% of the average citations per patent in the sample (6.48).

For the exploitative score in row 3 of Panel A, the difference between an inventor in the 25th percentile and that in the 75th percentile is 0.12. The economic effect is large given that the average exploitative score in our sample is 0.13. Finally, for the exploratory score in the last row of Panel A, the difference between an inventor in the 25th percentile and an inventor in the 75th percentile is 0.23, which is sizable given that the average exploratory score in the sample is 0.75.

5. 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 importance of inventors and firms would change in response to different inventor and firm characteristics.

5.1 Inventor centrality

In this subsection we examine whether inventors are more important in firms with more "key" inventors. We define whether an inventor is a key one based on her normalized degree of centrality in the universe of inventor co-authorship network. Following Hochberg, Ljungqvist, and Lu (2007), we calculate an inventor's normalized degree of centrality each year, which equals 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 one 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 equals $n/9$ in which 9 is the 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 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 *LnCitePat* as the dependent variable, and 320 firms and 146,083 inventors when using *Exploit* and *Explore* as the dependent variable.

We report the results of the AKM estimation in the subsample of firms with high centrality in Table 8. 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-squared 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 performance measured by patent counts and citations per patent, and about the same important in explaining innovation style. Specifically, 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 6.3 in the baseline regression to 8.1 in this subsample when using the patent counts as the dependent variable. The ratio increases from 1.44 to 1.5 when using *Exploit* as the dependent variable.

5.2 Inventor mobility

Team turnover is related to the retention of key human capital and affects firm performance (Cornelli, Simintzi, and Vig, 2016). We consider a subsample of firms with high mobility 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 firms' organizational capital should matter more. One the other hand, high inventor turnover could indicate that the firm imposes high standards and pursues talented inventors, which 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 high mobility subsample.[16] The subsample is comprised of 587 firms and 21,113 inventors when using patent counts and citations per patent as the dependent variable, and 443 firms and 17,156 inventors when using exploitative and explorative ratios as the dependent variable.

Table 9 reports the results using the AKM method in the subsample of firms with high inventor mobility. In each column, we report the explanatory powers of inventor fixed effects and firm fixed effects as well as the fraction of model R-squared explained by each set of variables in parentheses.

The results suggest that, inventor fixed effects are more important than firm fixed effects in explaining innovation performance and about the same in explaining innovation style. Specifically, in column 1, the ratio rises from 6.3 (in the baseline tests) to 9.3 in the regression with patent counts as the dependent variable; in column 2 in which patent citation is the dependent variable, this ratio increases rapidly from 8.6 (in the baseline tests) to 20.5. In the regressions with innovation style metrics as the dependent variable, this ratio stays similar, from 1.40 to 1.37 in column 3 and from 1.33 to 1.08 in column 4. All these results show that, in firms with high inventor mobility (i.e., high frequency in inventor turnovers), the relative contribution of inventor is more important in driving innovation performance, which supports our second arguments.

5.3 Innovation difficulty

The third dimension we consider when exploring heterogeneity is the subsample of firms in which innovation is more difficult to achieve. Because human capital tends to be more important in these firms, we postulate that inventors' human capital plays a more significant role in determining innovation output than firms' 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 four categories above. If

[16] 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.

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 SIC industry. 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 2,549 firms and 145,666 inventors in the regressions using patent counts and citations as the dependent variable, and 1,996 firms and 132,145 inventors in the regressions using *Exploit* and *Explore* as the dependent variable.

Table 10 presents the results that use 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-squared explained by each set of variables in parentheses. We also compute the ratio that captures the relative contribution of inventors and firms.

Compared with our baseline results, these ratios in the high-tech subsample are larger: in column 1 the relative importance ratio increases from 6.3 to 7.7; in column 2 this ratio raises from 8.6 to 11; 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 output than in an average industry. Table 10 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 that inventor fixed effects and firm fixed effects do not jointly explain innovation output.

6. Addressing inventors' endogenous moving

Because we rely on inventors' moving across firms to estimate the relative importance of firms' organizational capital and inventors' human capital, an important concern is that our results could be biased due to endogenous moving by inventors. For example, if an inventor moves because she expects changes in her innovation output, we could wrongly attribute the change in innovation output to the firm's organizational capital. Alternatively, if an inventor moves to a firm with similar performance, we could underestimate the contribution of firms' organizational capital to her innovation output. We perform five sets of tests to address this endogenous moving concern.

First, we compare the characteristics of firms to and from which inventors move. We also compare the characteristics of movers and stayers. These tests help gauge how serious the issue of inventors' endogenous moving is.

Panel A of Table 11 compares the characteristics of firms that inventors move to and move from. We find that firms that inventors leave are on average larger, older, with better operating performance as well as higher leverage. The results suggest that on average movers end up in a firm that appears to be smaller, younger and less profitable than their previous firms. To the extent that individuals seek better career opportunities in better firms when they voluntarily move, this finding suggests that inventor moving in our sample is more likely due to involuntary moving (e.g., being laid off or demoted). Panel B compares the characteristics of movers and stayers. Movers generally produce fewer patents than stayers, and movers' patents receive fewer citations than stayers. The results suggest that an average mover has worse innovation performance than a stayer. This finding again suggests that movers are unlikely to leave the current firm voluntarily and hence our setting is unlikely to be subject to endogenous inventor moving.

Second, assuming endogenous moving exists, we divide inventors into two groups. The first group includes inventors who "move up" in the sense that the firms they are joining have better past operating performance (i.e., ROA) than the ones they are leaving. The second group includes inventors who "move down", i.e., these inventors are joining firms that have worse past operating performance (i.e., ROA) than the ones they are leaving. We assume that an inventor who "moves up" is more likely to move voluntarily and who "moves down" is more likely to move involuntarily (and thus this type of moving is more exogenous). Therefore, we could overestimate firm fixed effects and hence underestimate the relative importance of inventors' human capital for "moving-up" inventors who are more subject to endogenous moving.[17]

In Table 12 Panels A and B, we report the results for "moving-down" and "moving-up" inventors, respectively. In Panel A we observe that the inventor is about 13 times as important as the firm in explaining patent counts in the group of inventors who move down. This finding helps alleviate endogenous moving concerns because it is unlikely that an inventor endogenously chooses to move down. In Panel B we see that the inventor is only 5 times as important as the

[17] When we divide the sample by inventor moves, we restrict the observations to the sample of movers who only move once (This is the sample we use for Table 10). This is because if the movers move multiple times, it is difficult to divide the sample according to their moving types (i.e., moving up vs. moving down, moving to similar firms vs. moving to different firms).

firm in explaining patent counts in the group of inventors who move up. The relative importance of the inventor to the firm in explaining citations is also smaller for the group of inventors who move up. Regarding innovation style, we do not observe a significant change in the relative importance of inventors' human capital and firms' organizational capital across the two groups of inventors. This observation indicates that inventors' moving decisions do not seem to be related to their own innovation style. Overall, the results suggest that while it seems that there is some overestimation of firm fixed effects for inventors who move up, endogenous moving by inventors does not appear to substantially alter our overall conclusion.

Third, another concern of our main results is that inventors could endogenously choose to move to firms with similar operating performance. In this case, we may underestimate firms' organizational capital and hence overestimate the relative importance of inventors' human capital. This is because we would attribute little innovation to firm fixed effects if inventors only move to firms with similar performance, given that it is less likely for us to observe a change in the inventors' patenting around the moves. However, firms could have contributed more to innovation had the inventors moved to the firms with different performance. To address this concern, in Panels C and D of Table 12, we divide the sample into two groups. The first group contains inventors who move to firms with similar performance. We define a similarly performing firm as those that are in the bottom quartile of the difference in ROA between the current firm and the previous firm. The other group contains inventors who move to firms with different performance, which is defined as the firms that are in the top quartile of the difference in ROA between the current firm and the previous firm.

The results reported in Panels C and D show that the inventor is 8.2 times as important as the firm in explaining patent counts when she moves between firms with similar performance, and the inventor is 7 times as important as the firm in explaining patent counts when she moves between firms with different performance. For citations per patent, exploratory score, and exploitative score, the inventor is also more important than the firm in terms of the explanatory power. The evidence once again suggests that endogenous moving could not completely account for our main results.

Fourth, even if inventors move to firms that appear very different in terms of operating performance, it is possible that these firms are operationally similar, which still causes the firm fixed effect to be underestimated. For example, if the firm to which an inventor is moving is in an industry different from her previous firm, the new firm is more likely to affect her patenting by providing vastly different access to resource and environment. In contrast, if the firm to which an inventor is moving is in the same industry as her previous firm, the new firm is less likely to affect her patenting. Hence, we are likely to underestimate firm fixed effects in the former situation than that in the latter case. To address this concern, we classify inventor moving into two groups. One group contains inventors who move within the same industry and the other group contains inventors who move across industries. Panels E and Panel F of Table 12 repeat the main tests that compare inventors who move within the same industry and move across industries, respectively. We find that inventor fixed effects are 6.1 times as important as firm fixed effects in explaining patent counts when inventors move across different industries, and are 12.9 times as important as firm fixed effects in explaining patent counts when inventors move within the same industry. In addition, in terms of citations per patent, exploitative score, and exploratory score, the inventor also appears to be much more important than the firm. As a matter of fact, the ratio of explanatory power of inventor fixed effects and firm fixed effects for inventors who move across different industries is very close to our baseline findings. Thus, the results suggest that inventors who move within the same industry and hence are mostly likely to be subject to endogenous moving have limited effects on our results.

Finally, if an inventor moves to a firm for which there is a better match between the firm and herself, then part of the change in her innovation output would come from the matching effect, rather than the firm-specific fixed effects in innovation. However, under the AKM method, we would attribute the change in innovation output to firm-specific fixed effects. In other words, this matching possibility may lead to an overestimation of firm fixed effects. We address the concern in two ways. First, we control for the previous firm and inventor performance in our regression, under the assumption that part of the matching effect would be reflected in firm and inventor quality. Second, we examine a subsample of movers who are less likely to be subject to the matching effect. Panel G of Table 12 considers inventors who have moved between firms with a change of patent output within 25% in 3 years.[18] While the maximum change is 25%, the mean change in the patent output is only 0.5 %, and the median change is 0%. Given such a small change in patent output, the matching issue bears very little effect in this subsample. The results are consistent with our main finding that inventors play a more important role than the

[18] The results are similar when we use 20% or 10% and when we use 1 year or 5 years.

firm in explaining innovation performance, and about the same role in explaining innovation style. As a matter of fact, the relative importance of inventors against the firm is even higher in this subsample that is supposedly subject to less matching problem, which indicates an overestimation of firm fixed effects caused by the matching problem.

Overall, the test results in this section suggest that our main findings that individual human capital plays a way more important role than firms' organizational capital in explaining innovation performance is not completely driven by inventors' endogenous moving decisions. We would also like to note that the matching issue is not specific to our method. It is present in any data that matches labor with capital. Finally, we acknowledge that one needs to be cautious in interpreting and generalizing our results because endogenous moving by inventors appears to play some role in our findings and we cannot completely rule it out.

7. Conclusion and Discussion

In this paper, we study persistence in inventor and firm innovation output, and investigate the role of inventor human capital in a firm's innovation output. We show innovation persistence at both the firm level and the inventor level. Furthermore, that time-invariant inventor fixed effects explain a majority of the heterogeneity in innovation performance in terms of patent counts and citations, while inventor fixed effects play a relatively less important role in explaining innovation style in terms of patent exploratory and exploitive scores. In the cross section, inventors contribute more to innovation output when they are better connected, are in firms with higher inventor mobility, and are in firms in which innovation is more difficult to achieve.

Our findings have important policy implications. Our results imply that human capital plays a crucial role in terms of generating high quality innovation for a firm, while firm organizational capital (such as corporate culture) affects the style of innovation output to a larger degree. Therefore, to build long-lasting growth, a firm should allocate more resources on attracting, training, and retaining talents. A firm also needs to be cautious with the culture it fosters and how it is going to affect the type of innovation generated by its employees.

Finally, we need to bear in mind three caveats when generalizing our results. First, similar to other studies that use movements of individuals (e.g., executives, venture capitalists, bankers, employees, etc.) as an identification strategy, our empirical setting is subject to the concern that inventors' moving could be endogenous. 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, although additional tests suggest that endogenous moving does not alter our conclusion. 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 inventor and the firm. 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 organization capital contribution to its long-term success.

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., R. Baghai, and K. 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.

Benner, M. J., and M. Tushman, 2002. Process management and technological innovation: A longitudinal study of the photography and paint industries. *Administrative Science Quarterly* 47, 676-706.

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

Bernstein, S., 2015. Does going public affect innovation? *Journal of Finance* 70(4), 1365-1403.

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

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., 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*.

Chemmanur, T., E. Loutskina, and X. Tian, 2014. Corporate venture capital, value creation, and innovation. *Review of Financial Studies* 27(8), 2434-2473.

Cho, C., J. Halford, S. Hsu, and L. Ng, 2016. Do managers matter for corporate innovation? *Journal of Corporate Finance* 36, 206-229.

Coase, Ronald H., 1937. The nature of the firm. *Economica* 4(16), 386-405.

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

Cohen L., K. Diether, and C. Malloy, 2013. Misvaluing Innovation. *Review of Financial Studies* 26 (3), 635-666.

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

Cornelli, F., E. Simintzi, and V. Vig, 2016. Team stability and performance: Evidence from private equity. Unpublished Working Paper.

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*.

Denis, D.J., 2012. The persistent puzzle of corporate capital structure: current challenges and new directions. *The Financial Review* 47, 631–643.

Denis, D.J., and S.B. McKeon, 2016. Persistent operating losses and corporate financial policies. *Unpublished working paper.*

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, P., X. Tian, and X. Yan, 2014. Financial development and innovation: Cross-country evidence. *Journal of Financial Economics* 112(1), 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.

Kaplan, S., B. Sensoy, and P. Strömberg, 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.

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.

Leary, M.T., and M.R. Roberts, 2005. Do firms rebalance their capital structures? *Journal of Finance* 60, 2575–2619.

Levinthal, Daniel A., and James G. March, 1993. The myopia of learning. *Strategic Management Journal* 14, 95-112.

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.

Nelson, R. R., 1991. Why do firms differ, and how does it matter? *Strategic Management Journal* 12(S2), 61-74.

Nelson, R. R., 1995. Recent evolutionary theorizing about economic change. *Journal of Economic Literature* 33(1), 48-90.

Porter, M. E., 1992. Capital choices: Changing the way America invests in industry. *Journal of Applied Corporate Finance* 5(2), 4-16.

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.

Smith, W. K., and M. L. Tushman, 2005. Managing strategic contradictions: A top management model for managing innovation streams, *Organization Science* 16, 522-536.

Solow, R. M., 1957. Technical change and the aggregate production function. *Review of Economics and Statistics*, 312-320.

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*.

Winter, S. G., 1984. Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization* 5(3-4), 287-320.

Winter, S. G., 2003. Understanding dynamic capabilities. *Strategic Management Journal* 24(10), 991-995.

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 unambiguously identified. We match all patents in subset A with the HBS database using patent number, resulting in a set of 6,270,074 matched inventor-patent observations denoted by set A.

(2) We divide subset B, consisting of all unmatched observations after step one into two groups: one group referred to as subset B1 that is comprised of inventor-patent observations in which each patent is filed by a single inventor; and the other group referred to as subset B2 that collects the remaining inventor-patent observations in which every patent is coauthored by multiple inventors. We then match all observations in subset B1 with set B by patent number, which leads to 22,555 inventor-patent-assignee observations corresponding to 11,461 inventor-patent observations in subset B1 as each patent may be possessed by several assignees. We then determine one assignee for each observation based on matched information in subset 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 subset 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 subset A using the 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 pin down all assignees for the 11,461 observations.

(3) For all observations in subset B2, they are patents filed by multiple inventors and belonged to various assignees. Using patent number, we join them with subset B to form all pairwise combinations and then select one assignee for each inventor-patent observation. 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 observations obtained in above three steps, our final matched sample consists of 6,531,703 inventor-patent observations 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 observations. If all patents filed by an inventor of a year belong to a single assignee, we assume that the inventor was hired by this particular assignee during patent filing years. 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 a 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 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.[19] Otherwise, we randomly pick one for him. This procedure leads about 4,251,546 inventor-year observations.

For our analysis, we augment our inventor-year sample in a time 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 assigning zero to patent counts and citations per patent.[20] This method comes with a caveat: that how we can accurately identify inventors' employer in gap years. Following the example above, 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 distinct companies, say, A and B. In other words, how do we decide the company affiliations of a

[19] Admittedly, this is an ad hoc assignment. To alleviate this concern, we repeat our analysis with 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. These alternative methods do not alter the nature of the results.

[20] We assign missing value to metrics that measures inventors' innovation style (exploratory ratio and exploitative ratio) in years with no patent filing.

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).[21] This procedure leads to 7,445,855 inventor-year observations in our augmented sample.

[21] We try other methods of assigning company affiliations to a mover: (1) assuming the inventor belongs to the old company A for all the transition years; (2) assuming the inventor belongs to the new company B for all the transition years. Also, we conduct our analysis in which we exclude all years with missing 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	
$LnExpnum_t$	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*;
$LnExpcit_t$	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*;
$Assets_t$	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 × #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 × Cash Flow ((#18+#14)/#8) plus 0.283 × Q ((#6+#199×#25-#60-#74)/#6) plus 3.189 × Leverage ((#9+#34)/(#9+#34+#216)) minus 39.368 × Dividends ((#21+#19)/#8) minus 1.315 × 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*.

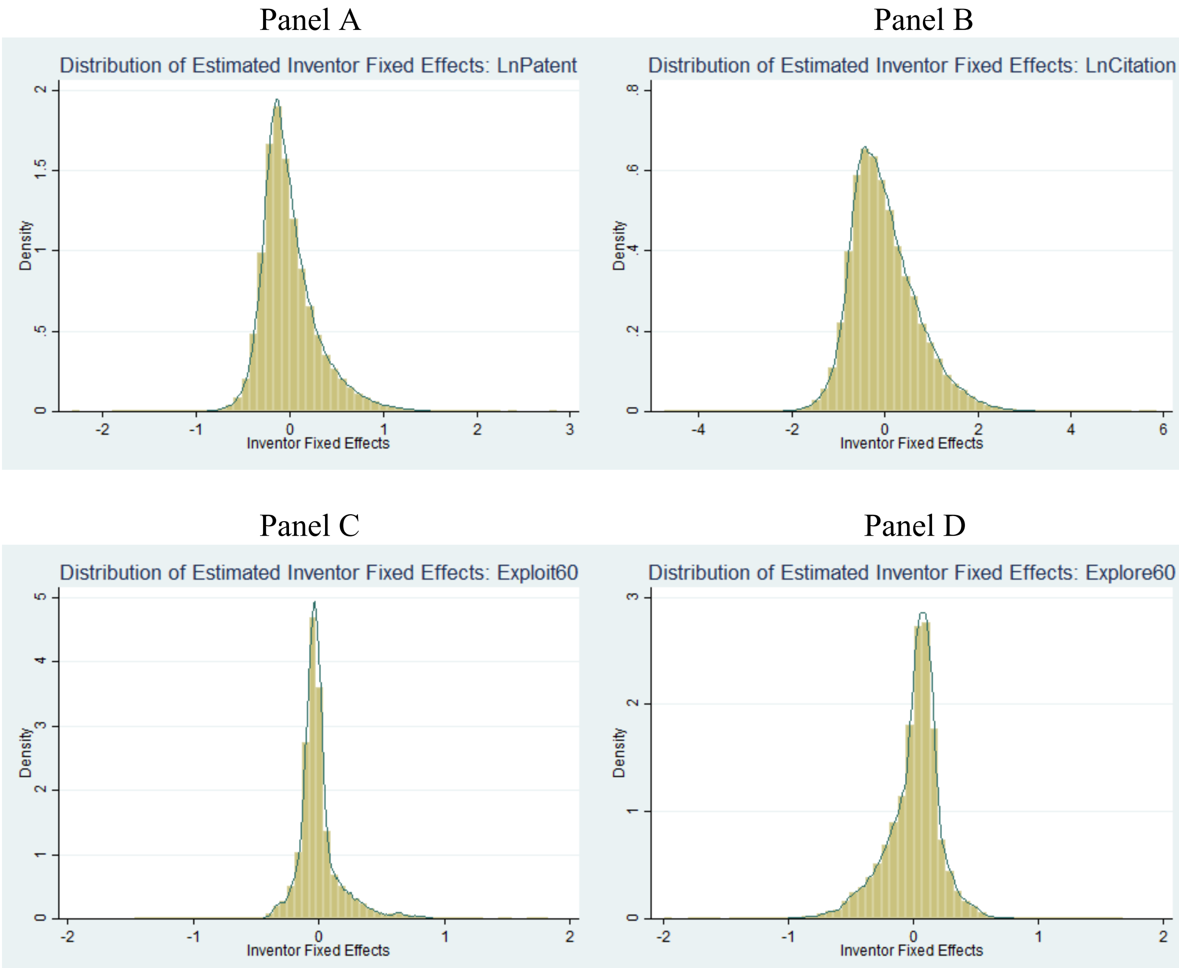

Figure 1. Distribution of estimated inventor fixed effects: connectedness sample. This figure displays the distribution of estimated inventor fixed effects from the AKM regression in the connectedness sample with four different dependent variables: the log of one plus the adjusted number of patents (Panel A), the log of one plus the adjusted number of citations per patent (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.

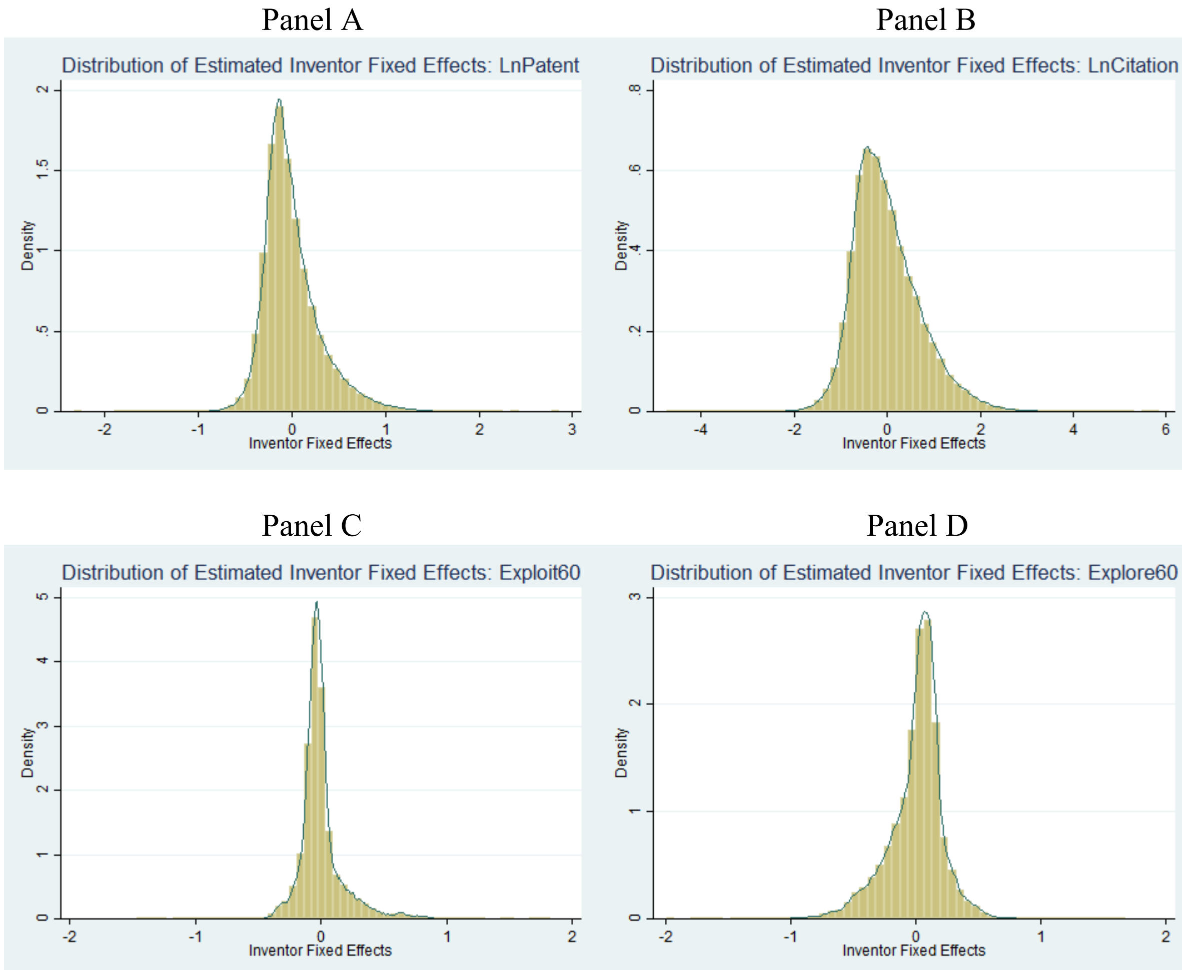

Figure 2. Distribution of estimated inventor fixed effects: largest group of the connected sample. This figure displays the distribution of estimated inventor fixed effects from the AKM regression in the largest connected sample with four different dependent variables: the log of one plus the adjusted number of patents (Panel A), the log of one plus the adjusted number of citations per patent (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 2003. A mover is defined as the inventor who switches employers. Panel A presents the employment information of inventors in our sample. Panel B presents the distribution of movers across firms.

Panel A: Number of movers out of all inventors

Mover	No. of firms in which inventors are employed	No. of inventors	Percentage
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
	7	2	0
	Subtotal	32,561	15.91
	Total	204,678	100

Panel B: Number of movers per firm

Mover Per Firm	Frequency	Percentage	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
31-50	268	4.68	87.68
51-100	253	4.42	92.10
>100	452	7.90	100
Total	5,722	100	

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 per patent as dependent variables in our baseline regression. Panel A presents the summary statistics of patent counts and citations per patent 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	0.910	0	1.484	0	1	1,246,951
Connectedness Sample	0.912	0	1.486	0	1	1,236,561
CitePat						
Full Sample	6.463	0	13.077	0	7.457	1,246,951
Connectedness Sample	6.476	0	13.091	0	7.481	1,236,561
Exploit						
Full Sample	0.126	0	0.307	0	0	555,592
Connectedness Sample	0.127	0	0.307	0	0	548,233
Explore						
Full Sample	0.745	1	0.401	0.5	1	555,592
Connectedness Sample	0.745	1	0.401	0.5	1	548,233

Panel B: Summary statistics for firms

Variable	Mean	Median	SD	25th	75th	N (firm-year)
Assets (million)						
Full Sample	7,120.888	691.71	21,416.62	110.87	4,084.982	46,177
Connectedness Sample	7,569.452	844.164	22,253.39	137.228	4,512	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	13.151	10.000	32.000	46,177
Connectedness Sample	22.386	22	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: Performance persistence measured by rankings

In this table, Panel A reports regressions where dependent variable is the dummy whether the performance of the inventor in terms of filed patent counts is in the top quantile (25%) every year (=1) or not (=0) and the independent variable is *Top_t* that the same dummy t years ago. Panel B reports the same regressions results as Panel A by using the dummy whether the performance of the inventor in terms of citations per patent is in the top quartile (25%) every year and the independent variable being the corresponding dummy t years ago. Panel C reports regressions where dependent variable is the dummy whether the performance of the firm in terms of filed patent number is in the top quantile (25%) every year (=1) or not (=0) and the independent variable is *Top_t* that the same dummy t years ago. Panel D reports the same regressions results as Panel C by using the dummy whether the performance of the firm in terms of citations per patent is in the top quartile (25%) every year and the independent variable being the corresponding dummy t years ago. Our specifications are linear models. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

Panel A: Inventor performance persistence (patent counts)

	(1)	(2)	(3)	(4)	(5)	(6)
Top_1	0.131***			0.122***		0.116***
	(0.001)			(0.001)		(0.001)
Top_2		0.085***		0.069***	0.080***	0.067***
		(0.001)		(0.001)	(0.001)	(0.001)
Top_3			0.029***		0.019***	0.012***
			(0.001)		(0.001)	(0.001)
Constant	-0.003	0.059***	0.097***	0.060***	0.097***	0.090***
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Adj. R^2	0.110	0.079	0.032	0.153	0.095	0.161
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,325,882	2,203,695	2,081,508	2,203,695	2,081,508	2,081,508

Panel B: Inventor performance persistence (citations per patent)

	(1)	(2)	(3)	(4)	(5)	(6)
Top_1	0.077***			0.078***		0.074***
	(0.001)			(0.001)		(0.001)
Top_2		0.050***		0.044***	0.051***	0.045***
		(0.001)		(0.001)	(0.001)	(0.001)
Top_3			0.008***		0.004***	0.001**
			(0.001)		(0.001)	(0.001)
Constant	0.015	0.038**	0.056***	0.035**	0.055***	0.050***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Adj. R^2	0.051	0.035	0.005	0.083	0.041	0.087
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,325,882	2,203,695	2,081,508	2,203,695	2,081,508	2,081,508

Panel C: Firm performance persistence (patent counts)

	(1)	(2)	(3)	(4)	(5)	(6)
Top_1	0.403***			0.339***		0.322***
	(0.006)			(0.007)		(0.007)
Top_2		0.262***		0.131***	0.239***	0.137***
		(0.006)		(0.007)	(0.007)	(0.007)
Top_3			0.147***		0.056***	0.012*
			(0.007)		(0.007)	(0.007)
Constant	-0.442***	-0.441***	-0.408***	-0.265***	-0.265***	-0.154***
	(0.036)	(0.041)	(0.047)	(0.039)	(0.046)	(0.044)
Adj. R^2	0.620	0.500	0.407	0.658	0.532	0.663
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,333	25,260	23,002	25,260	23,002	23,002

Panel D: Firm performance persistence (citations per patent)

	(1)	(2)	(3)	(4)	(5)	(6)
Top_1	0.072***			0.066***		0.057***
	(0.006)			(0.007)		(0.007)
Top_2		0.043***		0.039***	0.035***	0.031***
		(0.007)		(0.007)	(0.007)	(0.007)
Top_3			0.023***		0.021***	0.019***
			(0.007)		(0.007)	(0.007)
Constant	0.738***	0.710***	0.733***	0.660***	0.706***	0.668***
	(0.057)	(0.061)	(0.067)	(0.061)	(0.067)	(0.067)
Adj. R^2	0.068	0.047	0.034	0.081	0.047	0.074
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,333	25,260	23,002	25,260	23,002	23,002

Table 4: Persistence by number versus previous numbers

In this table, Panel A reports regressions where dependent variable is the number of filed patents for inventors at current year and the independent variable is *NPat_t* that measures the average number of filed patents in previous t years of the inventor. Panel B reports the same regressions results as Panel A by using the direct count of an inventor's citations per patent each year instead and the independent variable being the corresponding average counting in previous t years. Panel C reports regressions where dependent variable is the number of filed patents for firms at current year and the independent variable is *NPat_t* that measures the average number of filed patents in previous t year of the firm. Panel D reports the same regressions results as Panel C by using the direct count of citations per patent for a firm instead and the independent variable being the corresponding average counts in previous t years. All specifications are linear. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

Panel A: inventor sample with patent counts

	(1)	(2)	(3)	(4)
NPat_1	0.425***			
	(0.001)			
NPat_2		0.566***		
		(0.001)		
NPat_3			0.644***	
			(0.002)	
NPat_4				0.685***
				(0.003)
Constant	1.102***	0.954***	1.009***	0.885***
	(0.061)	(0.055)	(0.050)	(0.050)
Adj. R^2	0.112	0.284	0.375	0.419
Year FE	Yes	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes	Yes
Observations	1,163,050	734,823	520,405	391,544

Panel B: inventor sample with citations per patent

	(1)	(2)	(3)	(4)
NPat_1	0.126***			
	(0.001)			
NPat_2		0.202***		
		(0.002)		
NPat_3			0.243***	
			(0.002)	
NPat_4				0.270***
				(0.003)
Constant	1.711***	2.224***	2.766***	2.412***
	(0.624)	(0.527)	(0.475)	(0.439)
Adj. R^2	0.045	0.091	0.110	0.105
Year FE	Yes	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes	Yes
Observations	1,162,822	734,672	520,212	391,874

Panel C: firm sample with patent counts

	(1)	(2)	(3)	(4)
NPat_1	0.948***			
	(0.003)			
NPat_2		0.941***		
		(0.005)		
NPat_3			1.031***	
			(0.008)	
NPat_4				0.995***
				(0.011)
Constant	-39.656***	-52.371***	-35.562*	52.391
	(8.286)	(13.598)	(21.132)	(36.606)
Adj. R^2	0.949	0.934	0.927	0.909
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	14,392	8,790	5,894	4,159

Panel D: firm sample with citations per patent

	(1)	(2)	(3)	(4)
NPat_1	0.269***			
	(0.008)			
NPat_2		0.345***		
		(0.014)		
NPat_3			0.402***	
			(0.020)	
NPat_4				0.317***
				(0.031)
Constant	22.934***	17.069***	26.001***	25.692***
	(2.817)	(3.536)	(4.307)	(6.623)
Adj. R^2	0.284	0.320	0.279	0.265
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	14,392	8,790	5,894	4,159

Table 5: 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 using the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the natural logarithm of one plus the adjusted number of patents as the dependent variable and column 2 uses the natural logarithm of one plus the adjusted number of citations per patent as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 use the *Exploit* and *Explore* scores as the dependent variables, respectively (missing value is assigned if no patents filed by an inventor of a year). Contribution of inventor fixed effects $\frac{Cov\,(Y,\ Inventor\ FE)}{Var\,(Y)}$ captures the explanatory power of inventor fixed effects for innovation output (similarly for VC fixed effects). The percentages in parentheses present the fraction of model R-squared explained by each set of variables (see Graham, Li, and Qiu (2012), Ewens and Rhodes-Kropf (2015) for details). "Inv. FE / Firm FE" captures the ratio of the contribution of inventor fixed effects to the contribution of firm fixed effects in explaining innovation output. 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. Definitions of variables are defined in Appendix B. "# Firms" is the total number of firms in our sample. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.284 (53.08%)	0.266 (62.15%)	0.333 (55.50%)	0.318 (49.84%)
Contribution of Firm FE	0.045 (8.41%)	0.031 (7.24%)	0.231 (38.50%)	0.240 (37.62%)
Inventor FE / Firm FE	6.311	8.581	1.442	1.325
F-test on Fixed Effects				
Joint F-statistic	1.94***	1.87***	2.34***	2.58***
Inventor FE F-statistic	1.75***	1.67***	1.28***	1.34***
Firm FE F-statistic	2.91***	2.92***	3.79***	5.16***
Adj. R-squared after the Addition of Inventor and Firm Fixed Effects				
Control Variables	0.354	0.213	0.117	0.156
Adding Firm FE	0.373	0.237	0.337	0.387
Adding Inventor FE	0.438	0.308	0.379	0.431
Adding Both FE	0.442	0.313	0.393	0.450
# Movers	32,420	32,420	21,133	21,133
# Stayers	168,663	168,663	161,894	161,894
# Firms	4,294	4,294	3,249	3,249
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	1,231,352	1,231,352	547,923	547,923

Table 6: Different samples

This table reports regression results using the MDV method in Bertrand and Schoar (2003) to estimate both inventor and firm fixed effects of the mobility sample in Panel A and regression results using the AKM method to estimate both inventor and firm fixed effects of the largest group of connected sample in Panel B. Column 1 uses the natural logarithm of one plus the adjusted number of patents as the dependent variable and column 2 uses the natural logarithm of one plus the adjusted number of citations per patent as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 use the *Exploit* and *Explore* scores as the dependent variables respectively (missing value is assigned if no patents filed by an inventor of a year). Contribution of inventor fixed effects $\frac{Cov\,(Y,\ Inventor\ FE)}{Var\,(Y)}$ captures the explanatory power of inventor fixed effects in explaining innovation output (similarly for VC fixed effects). The percentages in parentheses present the fraction of model R-squared explained by each set of variables (see Graham, Li, and Qiu (2012), Ewens and Rhodes-Kropf (2015) for details). "Inv. FE / Firm FE" captures the ratio of the contribution of inventor fixed effects to the contribution of firm fixed effects in explaining innovation output. 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. Definitions of variables are defined in Appendix B. "# Firms" is the total number of firms in our 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

Dependent Variable	LnPatent (1)	LnCitePat (2)	Exploit (3)	Explore (4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-squared Explained)				
Contribution of Inv. FE	0.237 (49.17%)	0.206 (57.87%)	0.187 (40.22%)	0.171 (33.60%)
Contribution of Firm FE	0.050 (10.37%)	0.042 (11.80%)	0.255 (54.84%)	0.243 (47.74%)
Inventor FE / Firm FE	4.740	4.905	0.733	0.734
F-test on Fixed Effects				
Joint F-statistic	2.15***	1.96***	2.70***	2.81***
Inventor FE F-statistic	1.96***	1.72***	1.22***	1.27***
Firm FE F-statistic	2.15***	2.18***	4.90***	5.38***
Adj. R-squared	0.414	0.272	0.315	0.371
# Persons (#Movers)	32,420	32,420	21,133	21,133
# Firms	4,294	4,294	3,249	3,249
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	320,983	320,983	113,025	113,025

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

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.284 (53.08%)	0.266 (62.15%)	0.333 (55.50%)	0.318 (49.92%)
Contribution of Firm FE	0.045 (8.41%)	0.031 (7.24%)	0.231 (38.50%)	0.240 (37.68%)
Inventor FE / Firm FE	6.311	8.581	1.442	1.325
F-test on Fixed Effects				
Joint F-statistic	1.94***	1.87***	2.34***	2.58***
Inventor FE F-statistic	1.75***	1.68***	1.28***	1.34***
Firm FE F-statistic	2.97***	2.97***	3.88***	5.22***
Adj. R-squared	0.442	0.313	0.394	0.451
# Movers	32,310	32,310	20,991	20,991
# Stayers	168,328	168,328	161,346	161,346
# Firms	4,113	4,113	3,055	3,055
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	1,229,376	1,229,376	546,519	546,519

Table 7: 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 that 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 Inventors
LnPatent	-0.062	0.306	-0.196	0.137	201,083
LnCitePat	-0.112	0.727	-0.504	0.407	201,083
Exploit	-0.029	0.174	-0.084	0.035	183,027
Explore	0.037	0.223	-0.100	0.126	183,027

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

Dep. Variable	Median	SD	25th	75th	Number of Inventors
LnPatent	-0.062	0.306	-0.196	0.137	200,638
LnCitePat	-0.112	0.727	-0.503	0.407	200,638
Exploit	-0.029	0.174	-0.084	0.035	182,337
Explore	0.037	0.222	-0.100	0.125	182,337

Table 8: 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 of centrality. From 1970 to 2003, we calculate inventors' normalized degree centrality of a year, which equals 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 (see Hochberg, Ljungqvist, and Lu (2007) for more details 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 comprise a subsample of firms with high centrality. The estimation is implemented by using the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the natural logarithm of one plus the adjusted number of patents as the dependent variable and column 2 uses the natural logarithm of one plus the adjusted number of citations per patent as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 use the *Exploit* and *Explore* scores 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.

Dependent Variable	LnPatent (1)	LnCitePat (2)	Exploit (3)	Explore (4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.282 (50.90%)	0.260 (57.78%)	0.338 (56.52%)	0.319 (50.08%)
Contribution of Firm FE	0.035 (6.32%)	0.020 (4.44%)	0.224 (37.46%)	0.239 (37.52%)
Inventor FE / Firm FE	8.057	13.000	1.509	1.335
F-test on Fixed Effects				
Joint F-statistic	1.99***	1.92***	2.45***	2.72***
Inventor FE F-statistic	1.82***	1.74***	1.35***	1.40***
Firm FE F-statistic	9.99***	9.60***	18.43***	21.92***
Adj. R-squared	0.467	0.343	0.405	0.462
# Movers	17,657	17,657	10,751	10,751
# Stayers	149,132	149,132	135,109	135,109
# Firms	426	426	309	309
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	1,027,973	1,027,973	451,666	451,666

Table 9: 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 equals 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 using the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the natural logarithm of one plus the adjusted number of patents as the dependent variable and column 2 uses the natural logarithm of one plus the adjusted number of citations per patent as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 use 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.

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.307 (56.02%)	0.308 (69.84%)	0.318 (52.91%)	0.267 (41.40%)
Contribution of Firm FE	0.033 (6.02%)	0.015 (3.40%)	0.232 (38.60%)	0.247 (38.29%)
Inventor FE / Firm FE	9.303	20.533	1.371	1.081
F-test on Fixed Effects				
Joint F-statistic	1.72***	1.60***	1.44***	1.50***
Inventor FE F-statistic	1.54***	1.43***	0.87	1.05***
Firm FE F-statistic	1.86***	2.02***	4.64***	5.16***
Adj. R-squared	0.419	0.281	0.275	0.355
# Movers	3,018	3,018	2,167	2,167
# Stayers	18,095	18,095	14,989	14,989
# Firms	587	587	443	443
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	98,960	98,960	39,914	39,914

Table 10: 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 using the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the natural logarithm of one plus the adjusted number of patents as the dependent variable and column 2 uses the natural logarithm of one plus the adjusted number of citations per patent as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 use 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.

Dependent Variable	LnPatent (1)	LnCitePat (2)	Exploit (3)	Explore (4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.292 (53.19%)	0.266 (59.38%)	0.350 (58.04%)	0.334 (52.85%)
Contribution of Firm FE	0.038 (6.92%)	0.024 (5.36%)	0.216 (35.82%)	0.203 (32.12%)
Inventor FE / Firm FE	7.684	11.038	1.620	1.645
F-test on Fixed Effects				
Joint F-statistic	1.97***	1.90***	2.43***	2.53***
Inventor FE F-statistic	1.77***	1.70***	1.26***	1.34***
Firm FE F-statistic	3.17***	3.21***	3.94***	4.55***
Adj. R-squared	0.456	0.334	0.404	0.447
# Movers	19,767	19,767	13,555	13,555
# Stayers	125,899	125,899	118,590	118,590
# Firms	2,549	2,549	1,996	1,996
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	870,111	870,111	403,099	403,099

Table 11: Characteristics of firms and inventors

This table reports characteristics of firms that movers leave and move to, inventors that move and stay, and firms in and out of the sample. A mover is an inventor that switches firms during the sample period. Numbers reported are the mean across each subsample. Column 3 reports the differences in mean and their significance between the two samples. Panel A reports the differences in characteristics between firms which inventors move to and firms which inventors move from. Panel B reports the mean of innovation metrics for movers and stayers. "Moving From" is for firms that only have a mover that leaves the firm. "Moving To" is for firms that only have a mover that moves to the firm. Panel B reports the mean of innovation output of movers and stayers, and the difference in the mean of innovation output between movers and stayers. "Movers" are inventors that move at least once. "Stayers" are inventors that stay in the same company and never move. Definitions of variables are defined in the table of appendix. ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.

Panel A: Firms from and to which inventors move

	Moving To (1)	Moving From (2)	Difference (3)
Assets	24,222.86	27,195.5	-2,972.64***
RDAssets	0.064	0.057	0.007***
ROA	0.142	0.151	-0.009***
Leverage	0.215	0.231	-0.016***
PPEAssets	0.282	0.312	-0.030***
CapexAssets	0.070	0.077	-0.007***
TobinQ	2.058	1.663	0.385***
Founded Year	1971	1963	8.000***
Observations	3,615	2,384	

Panel B: Stayers vs. movers

	Movers (1)	Stayers (2)	Difference (3)
Patent	0.937	1.026	-0.089***
Citation	6.858	7.168	-0.310***
Exploit	0.072	0.114	-0.042***
Explore	0.812	0.757	0.055***
Observations	32,513	168,772	

Table 12: Different types of movers

This table reports the subsample analysis results using the AKM method to estimate both inventor and firm fixed effects in subsamples containing different types of movers. The connected group is constructed after restricting to this set of movers. The subsample of movers that move down starts with movers whose new firm's operating performance (defined as the average ROA of the firm in past three years) was higher than their previous firm's operating performance. The subset of movers that move up starts with movers whose new firm's operating performance was lower than their previous firm's operating performance. The subsample of movers that move to firms with similar operating performance considers the movers who moved to a firm that was in the bottom quartile of the difference between new versus previous firm performance. The subsample of movers that move to firms with difference operating performance considers the movers who moved to a firm that was in the top quartile of the difference between new versus previous firm performance. The different industry subsample considers movers who move across different industries. The same industry subsample considers movers who move within the same industries. Industry is defined based on 3-digit SIC codes. The subsample of movers that move with a minor change of output includes inventors who have moved between firms with a change of patent output within 25% in 3 years. Panel A - G correspond to the estimation with seven different subsamples described above. The connected group is constructed after restricting to this set of movers. If a mover moves more than once, it is excluded from the subsamples. The estimation is implemented by using the Stata command "felsdvreg" proposed by Cornelissen (2008). Column 1 uses the natural logarithm of one plus the adjusted number of patents as the dependent variable and column 2 uses the natural logarithm of one plus the adjusted number of citations per patent as the dependent variable (zero if no patents filed by an inventor of a year). Columns 3 and 4 use 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.

Panel A: Moving down

Dependent Variable	LnPatent (1)	LnCitePat (2)	Exploit (3)	Explore (4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.313 (56.29%)	0.290 (63.74%)	0.341 (55.63%)	0.336 (51.22%)
Contribution of Firm FE	0.024 (4.32%)	0.015 (3.30%)	0.227 (37.03%)	0.239 (36.43%)
Inventor FE / Firm FE	13.042	19.333	1.502	1.406
F-test on Fixed Effects				
Joint F-statistic	1.95***	1.90***	2.18***	2.49***
Inventor FE F-statistic	1.77***	1.72***	1.26***	1.34***
Firm FE F-statistic	2.09***	2.01***	1.94***	2.68***
Adj. R-squared	0.459	0.336	0.397	0.464
# Movers	5,218	5,218	2,898	2,898
# Stayers	159,038	159,038	142,125	142,125
# Firms	1,768	1,768	1,201	1,201
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	925,847	925,847	409,164	409,164

Panel B: Moving up

Dependent Variable	LnPatent (1)	LnCitePat (2)	Exploit (3)	Explore (4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.284 (51.26%)	0.270 (59.60%)	0.321 (52.45%)	0.305 (46.56%)
Contribution of Firm FE	0.051 (9.21%)	0.033 (7.28%)	0.246 (40.20%)	0.267 (40.76%)
Inventor FE / Firm FE	5.569	8.182	1.305	1.142
F-test on Fixed Effects				
Joint F-statistic	1.95***	1.90***	2.18***	2.50***
Inventor FE F-statistic	1.77***	1.72***	1.26***	1.34***
Firm FE F-statistic	2.32***	2.16***	1.82***	2.81***
Adj. R-squared	0.457	0.334	0.398	0.465
# Movers	5,599	5,599	2,936	2,936
# Stayers	159,129	159,129	141,933	141,933
# Firms	1,761	1,761	1,135	1,135
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	933,435	933,435	411,053	411,053

Panel C: Moving to firms with similar operating performance

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.294 (52.69%)	0.278 (60.43%)	0.451 (73.81%)	0.354 (54.13%)
Contribution of Firm FE	0.036 (6.45%)	0.022 (4.78%)	0.118 (19.31%)	0.223 (34.10%)
Inventor FE / Firm FE	8.167	12.636	3.822	1.587
F-test on Fixed Effects				
Joint F-statistic	1.99***	1.95***	2.28***	2.60***
Inventor FE F-statistic	1.80***	1.76***	1.29***	1.36***
Firm FE F-statistic	2.06***	1.99***	2.30***	3.42***
Adj. R-squared	0.463	0.345	0.401	0.467
# Movers	5,427	5,427	3,498	3,498
# Stayers	150,590	150,590	131,219	131,219
# Firms	1,307	1,307	892	892
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	895,237	895,237	387,661	387,661

Panel D: Moving to firms with different operating performance

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.295 (53.06%)	0.275 (60.44%)	0.323 (52.69%)	0.303 (46.26%)
Contribution of Firm FE	0.042 (7.55%)	0.030 (6.59%)	0.245 (39.97%)	0.271 (41.37%)
Inventor FE / Firm FE	7.024	9.167	1.318	1.118
F-test on Fixed Effects				
Joint F-statistic	1.95***	1.89***	2.18***	2.49***
Inventor FE F-statistic	1.77***	1.72***	1.25***	1.34***
Firm FE F-statistic	2.26***	2.16***	1.87***	2.79***
Adj. R-squared	0.459	0.335	0.398	0.464
# Movers	5,411	5,411	2,942	2,942
# Stayers	159,734	159,734	141,570	141,570
# Firms	1,841	1,841	1,202	1,202
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	930,188	930,188	409,775	409,775

Panel E: Moving across different industries

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.289 (52.64%)	0.273 (61.35%)	0.339 (55.57%)	0.326 (50.00%)
Contribution of Firm FE	0.047 (8.56%)	0.032 (7.19%)	0.227 (37.21%)	0.246 (37.73%)
Inventor FE / Firm FE	6.149	8.531	1.493	1.325
F-test on Fixed Effects				
Joint F-statistic	1.94***	1.88***	2.16***	2.46***
Inventor FE F-statistic	1.76***	1.70***	1.24***	1.33***
Firm FE F-statistic	2.52***	2.61***	2.45***	3.47***
Adj. R-squared	0.452	0.326	0.391	0.458
# Movers	14,443	14,443	7,473	7,473
# Stayers	165,613	165,613	152,929	152,929
# Firms	2,997	2,997	1,903	1,903
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	1,036,775	1,036,775	453,325	453,325

Panel F: Moving within the same industry

Dependent Variable	LnPatent	LnCitePat	Exploit	Explore
	(1)	(2)	(3)	(4)
Relative Importance of Inventor and Firm Fixed Effects (Percentage of R-sq. Explained)				
Contribution of Inv. FE	0.310 (55.96%)	0.285 (62.91%)	0.430 (69.47%)	0.450 (68.70%)
Contribution of Firm FE	0.024 (4.33%)	0.019 (4.19%)	0.147 (23.75%)	0.123 (18.78%)
Inventor FE / Firm FE	12.917	15.000	2.925	3.659
F-test on Fixed Effects				
Joint F-statistic	1.97***	1.92***	2.32***	2.56***
Inventor FE F-statistic	1.78***	1.73***	1.29***	1.35***
Firm FE F-statistic	2.24***	2.23***	2.99***	3.97***
Adj. R-squared	0.457	0.335	0.408	0.463
# Movers	11,107	11,107	7,033	7,033
# Stayers	156,462	156,462	138,270	138,270
# Firms	2,135	2,135	1,532	1,532
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	958,953	958,953	412,490	412,490

Panel G: Moving with a minor change of output

	(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	0.290 (51.69%)	0.269 (58.35%)	0.368 (59.84%)	0.332 (50.00%)
Contribution of Firm FE	0.029 (5.17%)	0.026 (5.64%)	0.197 (32.03%)	0.248 (37.35%)
Inventor FE / Firm FE	10.000	10.346	1.868	1.339
F-test on Fixed Effects				
Joint F-statistic	2.01***	1.97***	2.16***	2.54***
Inventor FE F-statistic	1.82***	1.79***	1.28***	1.36***
Firm FE F-statistic	2.02***	2.08***	1.15**	1.95***
Adj. R-squared	0.470	0.349	0.403	0.478
# Movers	2,938	2,938	910	910
# Stayers	115,790	115,790	101,990	101,990
# Firms	1,314	1,314	581	581
Year FE	Yes	Yes	Yes	Yes
Functional Area FE	Yes	Yes	Yes	Yes
Observations	700,903	700,903	292,440	292,440