Should Firms Avoid Relying on Key Employees? Evidence from Inventors*

Jin Wang
Lazaridis School of Business and Economics, Wilfrid Laurier University
75 University Avenue West, Waterloo, ON N2L 3C5
jwang@wlu.ca

First version: January 2023
This version: November 2024

Abstract

Using a large and unique dataset that tracks the career paths of inventors in U.S. public firms, we demonstrate the positive impact of inventor-base concentration on corporate innovation productivity in subsequent years. Moreover, we find that firms with a more concentrated inventor base tend to prioritize exploitation over exploration. These results are robust across various model specifications and sample selections. To understand the underlying mechanisms driving this positive impact, we evaluate two hypotheses: the *Human Capital Quality Hypothesis* and the *Employee Retention Hypothesis*. Our analysis reveals that inventor-base concentration improves the retention of both individual inventors and inventor teams. These findings highlight the importance of considering human capital strategies when evaluating a firm's innovation performance.

Keywords: inventor-base concentration, innovation productivity, exploitative innovation, exploratory innovation, inventor human capital quality, inventor retention

JEL classification: G34, O32, O34

* I thank the comments from Adam Jaffe and the attendees at the Financial Management Association European Conference (Aalborg, Denmark) and the Financial Management Association Annual Meetings (Dallas, TX). I acknowledge financial support from the Lazaridis Institute and the Social Sciences and Humanities Research Council of Canada (Grant Numbers: 435-2024-0363). All errors are my own.

Should Firms Avoid Relying on Key Employees? Evidence from Inventors

Abstract

Using a large and unique dataset that tracks the career paths of inventors in U.S. public firms, we demonstrate the positive impact of inventor-base concentration on corporate innovation productivity in subsequent years. Moreover, we find that firms with a more concentrated inventor base tend to prioritize exploitation over exploration. These results are robust across various model specifications and sample selections. To understand the underlying mechanisms driving this positive impact, we evaluate two hypotheses: the *Human Capital Quality Hypothesis* and the *Employee Retention Hypothesis*. Our analysis reveals that inventor-base concentration improves the retention of both individual inventors and inventor teams. These findings highlight the importance of considering human capital strategies when evaluating a firm's innovation performance.

Keywords: inventor-base concentration, innovation productivity, exploitative innovation, exploratory innovation, inventor human capital quality, inventor retention, operating performance

JEL classification: G34, O32, O34

1. Introduction

Human capital is crucial to both corporate innovation and the broader knowledge-based economy. While investing in human capital is essential for maintaining a firm's competitive advantage, a management dilemma arises: firms may struggle to fully capitalize on their human capital assets (Coff, 1999). Unlike physical assets, human capital is at risk due to voluntary turnover (Cascio, 1991; Steffy and Maurer, 1988; Diamond and Rajan, 2000). When employees leave, they take their valuable human and social capital with them, resulting in significant losses for the firm (Carnahan and Somaya, 2013; Dokko and Rosenkopf, 2010; Raffiee, 2017; Somaya et al., 2008; Jaravel et al., 2018; Wang and Zheng, 2022). As the traditional view suggests, overreliance on a small group of key employees is risky and detrimental to long-term success. A Forbes article aptly states: "Overall, key person dependency is bad for productivity and profits. It also stifles the growth of your other employees." [1]

Given the importance of human capital for competitive advantage and the risks of overdependence, it is surprising that existing research lacks evidence on how relying on key employees affects firm performance. This paper addresses this gap by examining the impact of inventor-base concentration on corporate innovation productivity.

To measure inventor-base concentration, we calculate each inventor's patent output over the past five years and compute the Herfindahl index based on the proportion of each inventor's output relative to the firm's total patent output (hereafter referred to as *Inventor concentration*). The index ranges from 0 to 1, with higher values indicating greater reliance on a concentrated group of inventors for innovation.

[1] See https://www.forbes.com/sites/forbestechcouncil/2021/12/28/the-risk-of-key-person-dependency-for-information/.

Our empirical analysis explores how inventor-base concentration influences firms' patenting output in years *t+1*, *t+2*, *and t+3*. The findings reveal a positive relationship between inventor-base concentration and both the quantity and citation-weighted quality of patents filed during these subsequent years. Specifically, a one-standard-deviation increase in *Inventor concentration* corresponds to a 7.22% rise in patent applications the following year. These results suggest that a strategy emphasizing a concentrated inventor base can significantly enhance innovation productivity.

Further, drawing on literature about the role of star inventors in corporate innovation (e.g., Tzabbar and Kehoe, 2014), we investigate the influence of inventor-base concentration on a firm's innovation style—specifically, the balance between exploration and exploitation. As anticipated, inventor-base concentration leads to a lower emphasis on exploration and a higher focus on exploitation in subsequent years.

To validate our findings, we conduct robustness checks. First, we confirm that the results hold after accounting for firms' technological innovation concentration and the size of their inventor base. Second, we show that the findings are not driven by firms with relatively small-scale innovation production, measured by patent output or inventor base size. Finally, using an instrumental variable approach based on the staggered adoption of the Inevitable Disclosure Doctrine (IDD) by U.S. state courts, we establish the causal effect of inventor-base concentration on corporate innovation productivity.

We also examine mechanisms through which inventor-base concentration enhances innovation productivity. Guided by human capital management literature, we propose two hypotheses: the *Human Capital Quality Hypothesis* and the *Employee Retention Hypothesis*. Detailed analyses at both the inventor and inventor-pair levels support the *Employee*

Retention Hypothesis, indicating that improved inventor retention, facilitated by a concentrated inventor base, is the primary mechanism driving enhanced innovation productivity.

Our paper contributes to two key streams of literature. First, it expands on research exploring the role of human capital in corporate innovation. Chemmanur et al. (2019, 2020) demonstrated the positive effects of top management quality on innovation output in public and private firms. Bhaskarabhatla et al. (2021) and Liu et al. (2017) highlighted the importance of inventors' human capital over firm capabilities in explaining innovation output variability. While these studies focus on the quantity of human capital, our work emphasizes the impact of its distribution across inventors. [2]

Second, this paper contributes to the literature on the benefits and costs of relying on key non-financial stakeholders. For example, Patatoukas (2012) found that customer-base concentration improves operating performance, while Irvine et al. (2016) and Hui et al. (2019) demonstrated that this benefit depends on the maturity of customer-supplier relationships. In contrast, we focus on inventors as key stakeholders and show that inventor-base concentration positively influences innovation outcomes.

Finally, our findings address a critical policy question: Should firms be protected from human capital loss due to voluntary turnover? The U.S. Federal Trade Commission (FTC) recently proposed a rule to ban noncompete agreements. By showing that reliance on key inventors—while increasing vulnerability to turnover—improves inventor retention and

[2] More broadly, our paper is related to the literature on firm-level determinants of corporate innovation. An incomplete list of the identified firm characteristics includes managerial overconfidence (Galasso and Simcoe, 2011), CEO personality (Sunder et al., 2017), CEO network connections (Faleye et al., 2014), compensation structures (Ederer and Manso, 2013; Baranchuk et al., 2014; Mao and Zhang, 2017; Chang et al., 2015), board independence (Balsmeier et al., 2017), firm-level anti-takeover provisions (Chemmanur and Tian, 2017), pro-diversity policies (Mayer et al, 2017), customer geographic proximity (Chu et al., 2019), and director job security (Hsu et al., 2024).

innovation productivity, our findings provide evidence that may inform policymakers like the FTC in shaping regulations on employer protections.

2. The Conceptual Framework

2.1 Reliance on key inventors and innovation productivity

The traditional view holds that relying on a small number of key employees is risky and detrimental to a firm's success. Unlike physical assets that a firm can control through ownership, leasing, or purchasing, human capital is embedded within employees and not under the firm's full control (Eisfeldt and Papanikolaou, 2013; Akins et al., 2020). These employees have the freedom to quit at will, taking their human and social capital with them (Carnahan and Somaya, 2013; Dokko and Rosenkopf, 2010; Raffiee, 2017; Somaya et al., 2008; Jaravel et al., 2018; Wang and Zheng, 2022). The negative impact of employee turnover on firm performance is even more severe when higher-performing employees leave (Shaw, Gupta, and Delery, 2005; Kwon and Rupp, 2013).

In firms that rely on a concentrated group of inventors to drive innovation, a significant portion of the firm's innovation capabilities are tied to the human capital of these key inventors (Groysberg and Lee, 2009; Paruchuri, 2010). The firm's innovation routines are integrated with the attributes of these key inventors, such as knowledge, information, ideas, skills, and relationships. If a key inventor leaves, it can disrupt the use of existing innovation routines and significantly undermine the firm's innovation capacities (Tzabbar and Kehoe, 2014). This reliance on key inventors for innovation production may pose a threat to the firm's innovation capabilities and negatively impact future innovation production (Aime et al., 2010).

While prior research highlights the risks of relying on key employees, it also demonstrates significant heterogeneity in both the quality of human capital and the impact of employees on firm performance (Wright et al., 2014). Some studies suggest that a small number of key employees with superior human capital often make disproportionately larger contributions than other employees (Fuller and Rothaermel, 2012; Liu, 2014; Tzabbar and Kehoe, 2014). To maximize its innovation capabilities, a firm may want to focus its investments in human capital on inventors with the greatest potential and rely on these key individuals for innovative activities.

Additionally, prior research demonstrates that as the productivity and career advancement opportunities for key inventors increase within a firm, they have fewer incentives to leave for a competitor (Hoisl, 2007) or to start their own business (Cassiman and Ueda, 2006; Kacperczyk, 2013; Sørensen and Sharkey, 2014). In firms that rely on key inventors for innovation, it is likely that these important individuals will be given better internal promotions, reducing the likelihood of voluntary turnover and minimizing the risk of losing valuable human capital.

Moreover, relying on key inventors inevitably incurs high replacement costs, thereby providing these individuals with increased job security. Manso (2011) argues that tolerance for failure is crucial in providing the optimal incentives for innovation. This view is supported by Acharya et al. (2014), who find that wrongful discharge laws, which protect employees from unjust termination, have a positive effect on innovation. Additionally, Acharya et al. (2013) find that more stringent labor dismissal laws drive innovation. Similarly, relying on key inventors for innovation production is likely to result in a higher tolerance for failure, motivating these individuals to take on risky projects and foster innovation.

Ultimately, it is an empirical question whether the benefits of a concentrated inventor base (potentially better quality in inventors' human capital and lower risk of voluntary or involuntary inventor turnover) outweigh its costs (greater loss due to key inventors' voluntary turnover).

2.2 Reliance on key inventors and exploitative/exploratory innovation

In its innovation production, a firm either searches for novel ideas closely related to its existing knowledge base (exploitation) or pursues new knowledge beyond its existing expertise (exploration). The choice between exploitation and exploration critically hinges on the inventors' expertise (Nelson & Winter, 1982).

With their central role in the firm's innovative activities, key inventors accumulate a disproportionately greater amount of expertise over time than other inventors. The tacit knowledge they hold is highly valuable, constituting a significant portion of the firm's entire knowledge base, yet it remains difficult for other inventors to observe and use. These key inventors can exert crucial influence on whether and how to utilize the firm's existing knowledge (Grant, 1996; Kogut and Zander, 1992).

From the standpoint of a firm that relies on a concentrated inventor base for innovation, the most effective way to utilize the expertise of these key inventors is to pursue innovation closely related to their areas of expertise. Since the accumulation of knowledge and skills is costly and time-consuming, key inventors are likely to leverage their past expertise. Pursuing less exploratory and more exploitative innovation provides both the firm and these key inventors with greater confidence and more consistent success in innovation production. This aligns with the findings of Tzabbar and Kehoe (2014), who discovered that exploration in innovation production increases after the departure of prolific inventors.

Based on this reasoning, we expect that a firm relying on a more concentrated inventor base in innovation production will pursue less exploration and more exploitation.

3. Measures of Inventor-base Concentration and Sample Formation

3.1. Measures of inventor-base concentration

To measure a firm's reliance on key inventors, we calculate the Herfindahl index based on each inventor's share of the firm's total patenting output during the five-year period up to year t. The share of an inventor's output is determined by dividing the number of patents filed by that inventor during the period by the total number of patents filed by all inventors in the firm during the same period. If a patent is filed by multiple inventors as collaborators, each inventor is considered to have contributed $1/n$ patents. The Herfindahl index ranges from 0 to 1, with a higher value indicating a greater reliance on key inventors in the firm's innovation production.

In robustness tests, we use an alternative measure of inventor-base concentration by calculating the proportion of patents filed by the most prolific inventor within the firm over the past five years relative to the total number of patents filed by the firm during the same period.

3.2. Sample formation

The information regarding patents is obtained from the United States Patent and Trademark Office's (USPTO) PatentsView database. This database contains detailed information about each patent granted between 1976 and 2021, including the date of application, patent citations, technology classes (classified using the Cooperative Patent Classification), the list of assignees (typically firms or their subsidiaries where the research was conducted), and the list of inventors. Additionally, the PatentsView database provides a unique identifier for each assignee and inventor, allowing us to track inventor-firm employment relationships over time using this information and the patent data.

We match the patents and patent assignees with U.S. public firms using the database provided by Stoffman, Yavuz, and Woeppel (2022), referred to as the SYW database. This database establishes the connection between patents and CRSP firms for patents granted from 1926 to 2021. By using the SYW database, we link the patents in the PatentsView database to the U.S. public firms that filed the patents and the inventors in the PatentsView database to the U.S. public firms where they work. Financial information about the U.S. public firms is obtained from Compustat.

Our sample consists of firm-year observations from both Compustat and CRSP for U.S.-based firms with common shares traded on NYSE, NASDAQ, or AMEX, and a history of filing at least one patent with the USPTO in the past five years. To eliminate firms with minimal economic impact, we include only those with book assets of at least ten million dollars and positive net sales. To avoid financially distressed firms, we also include only those with a positive book value of equity. Additionally, we exclude utility firms (SIC codes 4900-4999) and financial firms (SIC codes 6000-6999), as their performance is heavily influenced by regulations.

Our sample period begins in 1980, the earliest year we can measure a firm’s patenting activity over a five-year period (for example, the first five-year period covered is 1976-1980).

The PatentsView and SYW databases have data available up to 2021 as of the writing of this paper. Given the typical lag between patent application and grant, patents applied for up to 2019 are unlikely to be truncated. We measure a firm's patenting activity over a three-year period after each sample year, so our last sample year is 2016. Our final sample consists of 62,360 firm-year observations from 5,924 innovative firms.

3.3. Sample overview

Table 1, Panel A, presents the summary statistics of the sample firms' characteristics. These variables serve as the explanatory variables in our main analysis. We find that the average firm in our sample has a book value of assets equal to $6.2 billion, a Tobin's Q value of 2.08, and an R&D expenses to total assets ratio of 0.07. Notably, the average firm has an inventor-base concentration value of 0.252.

Panel B presents the summary statistics of the sample firms' innovation output in the subsequent years. These variables are used as the dependent variables in our main analysis. On average, the firms in our sample produce 30 patents in the first year, 29 patents in the second year, and 30 patents in the third year following the sample year.

Panel C presents the summary statistics of the sample firms' innovation styles in subsequent years. These variables are used as dependent variables in our analysis of the impact of inventor-base concentration on innovation style. On average, approximately 15% of patents produced each year are exploitative, while 34% are exploratory.

Panel D presents the Pearson correlations among the independent variables listed in Panel A. No variable pairs exhibit exceptionally high correlations that would raise concerns about multicollinearity. Notably, inventor-base concentration shows a negative correlation with the scale of innovation production.

4. Inventor-base Concentration and Firm Innovation

4.1. The baseline model

To investigate the impact of a strategy that relies on a concentrated inventor base on a firm's innovation productivity, we run pooled OLS regressions using the following baseline model:

$$
\begin{aligned}
&Innovation\ Output_{f,t+k} \\
&\qquad = \beta_1 Inventor\ Concentration_{f,t} + Innovation\ Output_{f,t-4\ to\ t} \\
&\qquad + \beta_2 Firm\ Characteristics_{f,t} + Firm\ FE_f + Year\ FE_t \\
&\qquad + e_{f,t}. \qquad (1)
\end{aligned}
$$

The dependent variable, $Innovation\ Output_{f,t+k}$, is the measure of the firm *f*'s innovation output in year $t + k$ (k is 1, 2, or 3). We measure innovation output using two variables based on the firm's patenting activities: 1) the raw number of patents filed; and 2) the citation-weighted number of patents filed, which takes into account the quality of the innovation by considering the number of citations received over five years after the patent award date and the median citations received by patents in the same technological class-year cell over the same period. The sample period is limited to year 2011 (five years prior to the end of the full sample period) when using the citation-weighted number of patents as the performance measure.

To capture the marginal effect of *Inventor concentration*, we control for the firm's innovation output, $Innovation\ Output_{f,t-4\ to\ t}$, over the past five years. Depending on the dependent variable, $Innovation\ Output_{f,t-4\ to\ t}$ can be either the number of patents *(# patents*) or the citation-weighted number of patents (# *citations*). # *patents* (or # *citations*) represents the number (or citation-weighted number) of patents filed by the firm over the past five-year period up to year *t*. Given that innovation capacities tend to persist over time, we expect these variables to have a positive coefficient.

We also control for a set of firm attributes, $Firm\ Characteristics_{f,t}$, that are known to impact innovation output. These variables are measured at the end of year *t* and constructed as follows:

- *Total Assets*: We use the book value of assets as a proxy for firm size, as larger firms are typically associated with greater levels of innovation output. Thus, we expect this variable to have a positive coefficient.
- *Tobin's Q*: We use the ratio of total market value of assets to book value of assets as a proxy for growth opportunities. Firms with greater growth opportunities are more likely to engage in innovation projects, so we expect this variable to have a positive coefficient.
- *ROA*: We use the ratio of income before extraordinary items to book value of assets as a measure of profitability. Since firms with higher profitability tend to have more resources available for innovation, we expect this variable to have a positive coefficient.

- *R&D*: We calculate the ratio of R&D expenses to total book value of assets as a measure of R&D intensity. Given that higher R&D input tends to be associated with greater innovation output, we expect this variable to have a positive coefficient.

To control for the effects of unobserved firm characteristics that may be associated with both inventor-base concentration and patenting output in subsequent years, we include firm fixed effects in our regression model. Additionally, to account for any shocks that occur across firms during the same time period, we also control for year fixed effects.

4.3. Regression results of the baseline model

The regression results of the baseline model are summarized in Table 2. In columns (1) to (3), the dependent variables are the raw number of patents filed in years *t+1*, *t+2*, and *t+3*, respectively. All coefficients for the control variables exhibit the expected signs and are statistically significant. Notably, the coefficients for *Inventor concentration* are positively significant across all three columns, indicating that a concentrated inventor base enhances innovation productivity. Economically, the results in column (1) suggest that a one standard deviation increase in *Inventor concentration* during the period from year *t-4* to year *t* leads to a 7.22% increase in patent applications in year *t+1*.

In columns (4) to (6), the dependent variables are the citation-weighted number of patents filed in years *t+1*, *t+2*, and *t+3*, respectively. The set of control variables remains consistent with columns (1) to (3), except that the number of citations replaces the number of patents. Once again, the coefficients for *Inventor concentration* are significantly positive across all three columns.

In summary, the panel regression results demonstrate a robust and significantly positive effect of a concentrated inventor base on innovation productivity in subsequent years.

4.4. Inventor-base concentration and innovation style

Motivated by the literature on the impact of star inventors on corporate innovation (e.g., Tzabbar and Kehoe, 2014), we investigate how inventor-base concentration influences a firm's innovation style, specifically the balance between exploitation and exploration.

To assess a firm's innovation in terms of exploration and exploitation, we define its existing expertise as the combination of its portfolio of patents filed and the citations received over the past five years. An exploratory patent utilizes knowledge beyond the firm's existing expertise, while an exploitative patent primarily builds on that existing expertise. Following Benner and Tushman (2002), a patent is considered "exploratory" if 80% or more of its citations are based on knowledge outside the firm's existing expertise, and "exploitative" if 80% or more of its citations refer to the firm's existing expertise. We calculate the ratios of exploratory and exploitative patents filed during a given time period relative to the total number of patents filed in the same period. These ratios reflect the extent to which the firm's innovation deviates from its accumulated knowledge base.

In Table 3, we examine how inventor-base concentration affects a firm's innovation style, using an augmented baseline model that includes the percentages of exploitative and exploratory patents filed in the past five years as additional explanatory variables. In columns (1) to (3), the dependent variables are the percentage of exploitative patents filed in years *t+1*, *t+2*, and *t+3*, respectively. We find that the coefficients for *Inventor concentration* are significantly positive, indicating that a concentrated inventor base positively influences exploitative innovations.

In columns (4) to (6), the dependent variables are the percentage of exploratory patents filed in years *t+1*, *t+2*, and *t+3*, respectively. We find that the coefficients for *Inventor concentration* are significantly negative, suggesting that a concentrated inventor base reduces exploratory innovations.

The results in Table 3 are consistent with the expectation that firms relying on key inventors tend to produce fewer exploratory innovations and more exploitative innovations.

5. Robustness tests

In this section, we conduct a series of robustness tests to confirm the positive impact of inventor-base concentration on innovation production in subsequent years.

5.1. Controlling for technology concentration and the scale of inventor base

To ensure that our results are not confounded by variables related to both inventor-base concentration and firm innovation, we augment the baseline model with additional control variables. First, key inventors often focus their innovative efforts within their areas of expertise (Tzabbar and Kehoe, 2014), making inventor-base concentration likely to correlate positively with technology concentration. This raises the possibility that inventor-base concentration may merely serve as a proxy for technology concentration. Second, the construction of the inventor-base concentration measure inherently leads to lower values as the number of inventors in a firm increases. Consequently, the results might reflect the scale of the inventor base rather than its concentration. Indeed, untabulated results reveal high correlations between inventor-base concentration, technology concentration (r = 0.663), and the number of inventors (r = -0.727).

Table 4 examines whether our results hold after controlling for these variables, which are defined as follows:

- *Technology concentration*: We calculate the Herfindahl index based on the proportion of patents filed by the firm across different patent classes over the past five years. The share of a technological class is calculated as the number of patents in that class divided by the firm's total patent count during the period. Since technology concentration could be associated with either higher innovation productivity (+) or a smaller scale of innovation production (-), we do not have a priori expectations about the sign of its coefficient.
- *# of inventors in past 5 years*: This variable measures the total number of distinct inventors who filed at least one patent in the previous five years. Given its likely positive relationship with the scale of innovation production, we expect a positive coefficient.

In Panel A, we replicate the baseline regressions from Table 2, adding *Technology concentration* as an explanatory variable. The coefficients for *Technology concentration* are significantly positive across all columns, indicating that firms with higher technology concentration achieve greater innovation productivity. Importantly, the coefficients for *Inventor concentration* remain significantly positive.

In Panel B, we include *# of inventors in past five years* as an additional explanatory variable. Consistent with expectations, its coefficients are positive and statistically significant across all columns. The coefficients for *Inventor concentration* remain significantly positive in all columns except column (6), where the coefficient is positive but becomes statistically insignificant.

In Panel C, we include both *Technology concentration* and # *of inventors in past five years* as additional explanatory variables. We show that the coefficients of *Inventor concentration* are significantly positive in all columns.

In Panel C, we incorporate both *Technology concentration* and # *of inventors in past five years* as additional explanatory variables. The coefficients for *Inventor concentration* remain significantly positive across all columns.

Overall, these results demonstrate that the positive relationship between inventor-base concentration and innovation productivity persists even after controlling for technology concentration and the size of the inventor base. This mitigates concerns that inventor-base concentration is merely a proxy for these variables.

5.2. Excluding firms with low innovation production

Given the negative correlation between inventor-base concentration and the scale of innovation output, another concern is that our results might be driven by firms with relatively small innovation production. To address this, we conduct robustness tests that exclude firms with low innovation output, measured either by patenting activity or the size of the inventor base. Table 5 presents the results.

In Panel A, we replicate the baseline regressions from Table 2, excluding firms whose patent output over the past five years is at or below the 25th percentile for their respective years. [3] The coefficients for *Inventor concentration* remain significantly positive across all columns.

In Panel B, we exclude firms where the number of inventors who filed at least one patent in the past five years is at or below the 25th percentile for their respective years. [4] Again, the coefficients for *Inventor concentration* remain significantly positive across all columns.

In summary, the results in Table 5 confirm that the positive impact of inventor-base concentration on innovation productivity is not driven solely by firms with low innovation production. These findings provide further evidence supporting the robustness of our main conclusions.

5.3. Instrumental variable regressions

Since our baseline models control for firm fixed effects, it is unlikely that our main findings are driven by the correlation between inventor-base concentration and any unobserved, time-invariant firm characteristics associated with innovation productivity. However, it is possible that inventor-base concentration is correlated with unobserved, time-variant firm characteristics that are also correlated with innovation productivity. To address this potential endogeneity issue, we use an instrumental variable approach in this subsection.

To construct the instrumental variable, we employ the staggered adoption of the Inevitable Disclosure Doctrine (IDD) by U.S. state courts, which prevents employees with knowledge of a firm's trade secrets from working for another firm. Using an instrumental

[3] For the full sample, the 25th percentile of the number of patents filed over the past five years is two.
[4] For the full sample, the 25th percentile of the number of inventors who filed at least one patent over the past five years is four.

variable based on the staggered recognition of the IDD by state courts is appealing for two reasons. First, state courts adopt the IDD to protect trade secrets for firms located in the state, reducing the risk of departing employees revealing these secrets to other firms. This protection is hence likely to impact the firm's incentive to adopt a human capital strategy that focuses on key inventors. [5] Second, the staggered adoption of the IDD is exogenous to firms, and the motivation behind the IDD is unrelated to fostering firm innovation.

Following Chen et al. (2021), we measure a firm's IDD protection based on its headquarters location. Specifically, we create a dummy variable that equals one if the firm's headquarters is located in a state that has adopted the IDD and zero otherwise. We then use this IDD dummy as an instrumental variable for inventor concentration to help establish the causal effect of inventor-base concentration on corporate innovation productivity. We obtain historical headquarters information from Bai et al. (2020) for the period before 1987, Compact Disclosure for 1987-2001, and Compustat for post-2001. The information about IDD adoption years in each state is from Klasa et al. (2018).

Table 6 presents the results of the instrumental variable regressions. Panel A shows the results of the first-stage regressions, where inventor concentration is the dependent variable, and firm characteristics, the IDD dummy, and firm dummies are the independent variables. Given that the value of the IDD dummy is entirely determined by the firm's identity and the year, and we control for firm fixed effects, we drop the year dummies to avoid multicollinearity.[6] Panel A shows that inventor-base concentration has significant negative

[5] On one hand, the adoption of the IDD reduces the risk of losing key inventors, thereby incentivizing a high inventor-base concentration. On the other hand, enhanced employer protection may reduce the firm's incentives to incur extra costs to retain key employees, potentially leading to a lower inventor-base concentration. Therefore, it is an empirical question whether the adoption of the IDD has a positive or negative impact on inventor-base concentration.

[6] Except for the rare cases where the firms changed the location of their headquarters over their history.

associations with past five-year patenting output, total book assets, and R&D intensity, and significant positive associations with Tobin's Q. Importantly, inventor-base concentration has a significant negative association with the IDD dummy, regardless of the past patenting output measure used.

Panel B shows the results of the second-stage regressions. Except for column (6), where the instrumented *Inventor concentration* has an insignificant coefficient, all coefficients for instrumented *Inventor concentration* remain significantly positive.

6. The Mechanisms

The evidence from the previous sections shows a causal relationship between inventor-base concentration and corporate innovation productivity in subsequent years. These results suggest that the benefits of a concentrated inventor base as a human capital strategy surpass the risks of losing key inventors. This section aims to gain insights into how inventor-base concentration promotes corporate innovation through in-depth analysis at the levels of individual inventors and inventor pairs.

The discussions in Section 2.1 suggest two possible mechanisms through which inventor-base concentration positively impacts innovation productivity:

- *Human Capital Quality Hypothesis*: Firms with a greater concentration of inventors tend to invest more heavily in the human capital of those inventors with the most potential, relying on these key individuals for innovative activities. To test this hypothesis, we analyze the relationship between inventor-base concentration and the human capital quality of individual inventors within a firm. If this hypothesis is correct, we expect to see a positive correlation between inventor-base concentration and inventor human capital quality.

- *Employee Retention Hypothesis*: Firms with a greater concentration of inventors are expected to offer better internal promotions to key inventors, thereby reducing voluntary turnover. This focus on key individuals also increases their job security, encouraging them to take risks in pursuit of innovation. If this hypothesis holds, we anticipate a positive correlation between inventor-base concentration and the inventor retention rate in the following years.

While the two hypotheses make different predictions, it is noteworthy that they are not mutually exclusive – that is, both employee quality and employee retention could be contributing factors to the positive impact of inventor-base concentration on innovation productivity.

6.1. The Human Capital Quality Hypothesis

To test the *Human Capital Quality Hypothesis*, we estimate the following pooled OLS regression:

$$Inventor\ human\ captial\ quality_{i,f,t} = \beta_1 Inventor\ concentration_{f,t} + \beta_2 Firm\ characteristics_{f,t} + \beta_3 Inventor\ characteristics_{i,t} + Inventor\ FE_i + Year\ FE_t + e_{i,f,t}. \quad (2)$$

The sample consists of all inventors who filed at least one patent within the sample firms over the five-year period from year *t-4* to year *t*. The dependent variable,

$Inventor\ human\ captial\ quality_{i,f,t}$, assesses the quality of inventor human capital within the firm at year *t*. To proxy for this quality, we use three variables: the inventor's total patent count up to year t, their citation-weighted patent count up to year t, and a binary variable indicating whether they were a "star inventor" in year *t* (defined as having a citation-weighted patent count among the top 5% of all inventors in year *t*).

The control variables comprise both firm and inventor characteristics, and include fixed effects for both the inventor and the year. The set of firm-level control variables is identical to those in Equation (1). The set of inventor-level control variables, measured at the end of year t, is defined as follows:

- *Inventor seniority*: the number of years between the first patent filed by inventor *i* over their entire career and year *t*.

Table 7 presents the regression results. In column (1), the dependent variable is the inventor's total patent count up to year *t*. We show that the inventor's total patent count is positively associated with *Inventor concentration*, at a significance level of 1%. In column (2), the dependent variable is the inventor's citation-weighted patent count up to year *t*. We find that the inventor's total patent count is negatively associated with *Inventor concentration*, at a significance level of 5%. In column (3), the dependent variable is the dummy variable for star inventor. The results indicate that the probability of being a star inventor is not associated with *Inventor concentration*. Overall, these mixed results do not support the *Human Capital Quality Hypothesis*.

6.2. The Employee Retention Hypothesis

To test the *Employee Retention Hypothesis*, we examine how inventor-base concentration affects the decision of inventors to remain with the firm in subsequent years, at both the individual inventor and the inventor-team level.

To perform an inventor-level analysis, we estimate the following pooled OLS regression, drawn from inventors working in our sample firms over our sample period:

$$Inventor\ stay_{i,f,t+k} = \beta_1 Inventor\ concentration_{f,t} + \beta_2 Firm\ characteristics_{f,t} + \beta_3 Inventor\ chracteristics_{i,t} + Inventor\ FE_i + Year\ FE_t + e_{i,f,t}. \quad (3)$$

The sample includes all inventors who filed at least one patent in the sample firms during the five-year period, from year *t-4* to year *t*. The dependent variable, $Inventor\ stay_{i,f,t+k}$, is a dummy variable that indicates whether the inventor remained with firm *f* in year *t + 1*, *t+2*, or *t+3*. We use the procedure described in Appendix A, to obtain information about an inventor's employer identity.

The control variables comprise both firm and inventor characteristics, and include fixed effects for both the inventor and the year. The set of firm-level control variables is identical to those in Equation (1). The set of inventor-level control variables, measured at the end of year *t*, is defined as follows:

- *# patents by inventor*: the logarithm of one plus the number of patents filed by inventor *i* up to year *t*.
- *Inventor tenure*: the number of years between the first patent filed by inventor *i* in firm *f* and year *t*.

Table 8 presents the regression results. We show that inventors are less likely to stay in the firm in years *t+1*, *t+2*, or *t+3* after they have been more productive over the past five-year period, suggesting that productive inventors have better external career opportunities. We also show that inventors are less likely to stay in the firm in years *t+1*, *t+2*, or *t+3* as their tenure in the firm increases. More importantly, the coefficients for *Inventor concentration* are significantly positive for all three columns, suggesting that having a concentrated inventor base positively impacts inventor retention.

To perform the inventor team-level test, we estimate the following pooled OLS regression using a sample of inventor pairs in our sample firms:

$$Collaborator\ stay_{i,j,f,t+k} = \beta_1 Inventor\ concentration_{f,t} + \beta_2 Firm\ characteristics_{f,t} + + \beta_3 Collaborator\ chracteristics_{j,t} + \beta_4 Inventor\ chracteristics_{i,t} + \beta_5 Inventor - pair\ chracteristics_{i,j,t} + Inventor\ pair\ FE_{i,j} + Year\ FE_t + e_{i,j,f,t}. \quad (4)$$

The sample includes all collaborator-inventor pairs, where the collaborator is the "significant collaborator" of the inventor and both the collaborator and the inventor filed at least one patent during the five-year period from year *t-4* to year *t*. A collaborator is defined as significant if they are the inventor's most frequent co-filer of patents in the preceding five years. The dependent variable, $Collaborator\ stay_{i,j,f,t}$, is a dummy variable that indicates whether the collaborator *j* of inventor *i* remained with firm *f* in years *t + 1*, *t+2*, or *t + 3*.

The control variables consist of firm, collaborator, inventor, and inventor-pair characteristics, with fixed effects for both the inventor pair and the year. The sets of firm-level and inventor-level variables are the same as those in Equation (2), while the inventor-pair level variables, $Inventor - pair\ chracteristics_{i,t}$, are measured at the end of year *t* and includes the following:

- *Frequency of collaborations*: the logarithm of one plus the number of collaborations between the collaborator and the inventor.
- *Inventor distance*: the logarithm of the geographical distance between collaborator *j* and inventor *i*.

Table 9 presents the results, which are consistent with prior findings. Notably, the likelihood of a collaborator's retention increases if the frequency of collaboration with the inventor is higher or the geographical distance between them is shorter. More importantly, the coefficients for *Inventor concentration* are significantly positive in all three columns, indicating that the strategy of building a concentrated inventor base has a positive impact on inventor-team retention.

To conclude, the findings from Tables 8 and 9 indicate that a concentrated inventor base has a positive impact on both inventor and inventor-team retention, supporting the *Employee Retention Hypothesis*.

The results in this section suggest that high inventor retention is a contributing factor to the positive impact of inventor-base concentration on corporate innovation, while inventor human capital quality does not appear to be a factor.

7. Alternative Measure of Inventor-base Concentration

Thus far, we have used the Herfindahl index, based on the share of patents across all inventors, as a measure of inventor-base concentration. To verify the robustness of our findings, we also use an alternative measure: the proportion of patents filed by the most prolific inventor within the firm over the past five years relative to the total number of patents filed by the firm during the same period. Table 10 replicates the previous analyses using this alternative measure and produces results that are qualitatively comparable to those obtained using the Herfindahl index.

8. Conclusions

Using a large and unique dataset that tracks the career paths of inventors in U.S. public firms, we find a positive impact of inventor-base concentration on corporate innovation productivity in subsequent years. Additionally, we examine the influence of inventor-base concentration on innovation styles and show that firms with a more concentrated inventor base tend to produce fewer innovations that deviate from their existing knowledge base, i.e., fewer exploratory innovations.

To understand the driving forces behind this positive impact, we test two hypotheses: the *Human Capital Quality Hypothesis* and the *Employee Retention Hypothesis*. While we do not find evidence that firms with a more concentrated inventor base have inventors with better human capital quality, we demonstrate that inventor-base concentration enhances the retention of both individual inventors and inventor teams.

These results underscore the importance of understanding human capital strategies when evaluating a firm's innovation performance.

Appendix A: Tracking an inventor's patenting career

To determine an inventor's employer(s) throughout her patenting career, we rely on inventor and assignee information in the PatentsView database (https://www.patentsview.org) and patent-PERMCO (i.e., the firm identifier in the CRSP database) link in the SYW database. We proceed in the following steps.

Step 1

Using the PatentsView database, we first identify all inventor-year pairs in which an inventor applied for at least one patent in that year. For each inventor-year pair, we then obtain assignees associated with all patents of the inventor. If there is only one assignee for all her patents filed in that year, the inventor's employer for that year is unambiguously identified. If there are multiple assignees for her patents filed in that year, the assignee with which the inventor filed the greatest number of patents in the year is identified as her employer.

Step 2

The process from Step 1 divides inventor-year pairs into two sets: those associated with a unique assignee (*UA*) and those associated with multiple assignees (*MA*, representing 13% of the sample). We determine the employer of an inventor-year pair in *MA* using the matched information in *UA*. Specifically, we match an assignee to an inventor-year pair in MA if the inventor has been matched to the same assignee in *UA* for the following year: $t-1$, $t+1$, $t-2$, $t+2$, $t-3$, and $t+3$, where t is the year of the inventor-year pair in *UA*. If we cannot determine an assignee for an inventor-year pair based on the matched information in *UA*, we randomly pick one of the assignees. The above process results in matched inventor-assignee-year observations for years in which an inventor applied for patents.

Step 3

We augment the inventor-assignee-year (*I-A-Y*) sample from Step 2 by filling gaps in which an inventor is not matched to an assignee as follows. If both *I-A-Y1* and *I-A-Y2* are observations in the sample and there are no other observations of inventor *I* between year *Y1* and *Y2*, then we assume inventor *I*'s employer is *A* during the period from *Y1* to *Y2*. If both *I-A1-Y1* and *I-A2-Y2* are observations in the sample and there are no other observations of inventor *I* between year *Y1* and *Y2*, then we assume inventor *I*'s employer is *A1* during the period from *Y1* to *Ym* and *A2* during the period from *Ym+1* to *Y2*, where $Ym = int\ (Y1 + (Y2 - Y1) / 2)$.

By the end of Step 3, we obtain inventor-assignee-year information on each inventor's active career that spans the year of her first patent application and the year of her last patent application in the PatentsView database.

Step 4

Using the patent-PERMCO link in the SYW database, we further match inventor-assignee-year observations to U.S. public firms. Specifically, we first merge patent-PERMCO pairs in SYW and patent-assignee pairs in PatentsView by patent number, and keep only those patent-PERMCO pairs in which a patent has a solo assignee. We then merge the resulting assignee-

PERMCO pairs with the inventor-assignee-year sample from Step 3 by patent number and obtain the sample of inventor-PERMCO-year observations.

<u>*Step 5*</u>

The inventor-PERMCO-year sample from Step 4 can be divided into two sets: those inventor-year pairs associated with a unique public firm (*UP*) and those inventor-year pairs associated with multiple public firms (*MP*). For inventor-PERMCO-year observations in *UP*, the public firm is identified as the employer of the inventor for the year. For inventor-PERMCO-year observations in *MP*, we use information on the starting and ending dates of firm names provided by CRSP to help filter out firms if the date range of the matched firm name does not cover the focal year. For those inventor-year pairs that are still associated with multiple firm names, we manually check and pick the most likely match.

By the end of Step 5, we obtain inventor-PERMCO-year information on each inventor's active career that spans the year of her first patent application and the year of her last patent application in the PatentsView database.

Appendix B: Variable definitions

Measures of inventor-base concentration	
Inventor concentration	The Herfindahl index based on the share of each inventor's patenting output over the five-year period up to year *t*. The share of an inventor's output is calculated as the number of patents filed by the inventor during the period scaled by the total number of patents filed by all inventors in the firm during the same period. For a patent filed by *n* inventors as collaborators, each inventor is deemed to have produced *1/n* patents.
Top inventor's share	The proportion of the most prolific inventor's output in the firm's total patenting output over the five-year period up to year *t*. The most productive inventor is the inventor who filed the greatest number of patents during the period. For a patent filed by *n* inventors as collaborators, each inventor is deemed to have produced *1/n* patents.
Measures of firm innovation performance	
# Patents	The numbers of granted patents filed by the firm in year *t+1*, *t+2*, *and t+3*, respectively.
# Citations	The number of citations received over the five-year period after patent award by patents filed in year *t+1*, *t+2*, and *t+3*, respectively. For each patent, the number of citations is scaled by the median number of citations received by patents in the same class-year cell.
% Exploitative patents	The percentage of exploitative patents among all patents filed by the firm in year *t+1*, *t+2*, and *t+3*, respectively. The firm's existing expertise is defined as the combination of its portfolio of patents filed over the past five years and the citations made by those patents. Following Benner and Tushman (2002), a patent is categorized as "exploitative" if 80% or more of its citations are made to the firm's existing expertise.
% Exploratory patents	The percentage of exploratory patents among all patents filed by the firm in year *t+1*, *t+2*, and *t+3*, respectively. Following Benner and Tushman (2002), a patent is categorized as "exploratory" if 20% or fewer of its citations are made to the firm's existing expertise.
Firm characteristics for predicating innovation output	
# Patents in past 5 years	The number of patents filed by the firm over the five-year period up to year *t*.
# Citations in past 5 years	The total number of citations received during the five-year period starting the grant date of the firm's patents filed over the five-year period up to year *t*. For each patent, the number of citations is scaled by the median number of citations received by patents in the same class-year cell.
Book assets	The book value of total assets in the firm (in 2021 dollar) at the end of year *t*.
Tobin's Q	The ratio of market value of total assets to the book value of total assets at the end of year *t*.
Profitability	The ratio of operating income before depreciations in year *t* to the book value of total assets at the end of year *t*.
R&D intensity	The ratio of R&D expenditures in year *t* to the book value of total assets at the end of year *t*.

Inventor and inventor-pair characteristics

# Patents by inventor	The number of patents filed by the focal inventor over their entire career up to year *t*.
# Citations by inventor	The total number of citations received over the five-year period after patent award by patents filed by the focal inventor over their entire career up to year *t*. For each patent, the number of citations is scaled by the median number of citations received by patents in the same class-year cell.
Star inventor	A dummy that indicates whether an inventor is a star in year *t*. A star inventor is among the top 5% inventors in terms of the scaled number of citations over their entire career up to year *t*.
Inventor stay	For an inventor who filed at least one patent with the firm over the five-year period ended in year *t*, the dummy indicates whether the inventor remains with the firm in year $t + k$ (k is 1, 2, or 3).
Inventor seniority	The number of years between the year when the focal inventor filed the first patents over their entire career and year *t*.
Inventor tenure	The number of years between the first year when the focal inventor works for the firm and year *t*.
# Patents by significant collaborator	The number of patents filed by the focal inventor's significant collaborator. The significant inventor is the inventor who files the greatest number of patents with the focal inventor over the five-year period up to year *t*.
Significant collaborator tenure	The number of years between the first year when the focal inventor's significant collaborator works for the firm and year *t*.
Frequency of collaboration	The number of patents filed by both the focal inventor and the significant collaborator over the five-year period up to year *t*.
Inventor-collaborator distance	The geographical distance between the locations of the focal inventor and the significant collaborator.

References:

Aime, Federico, Scott Johnson, Jason W. Ridge, and Aaron D. Hill, 2010. The routine may be stable but the advantage is not: Competitive implications of key employee mobility. *Strategic Management Journal*, 31: 75-87.

Akins, Brian, David De Angelis, and MacLean Gaulin, 2020. Debt contracting on management. *Journal of Finance*, 75: 2095-2137.

Acharya, Viral V, Ramin P. Baghai, and Krishnamurthy V. Subramanian, 2013. Labor laws and innovation. *Journal of Law and Economics*, 56: 997–1037.

Acharya, Viral V., Ramin P. Baghai, and Krishnamurthy V. Subramanian, 2014. Wrongful Discharge Laws and Innovation. *The Review of Financial Studies*, 27: 301-346.

Bai Jianqiu, Douglas Fairhurst, and Matthew Serfling, 2020. Employment protection, investment, and firm growth. *Review of Financial Studies*, 33:644–688.

Baranchuk, Nina, Robert Kieschnick, and Rabih Moussawi, 2014. Motivating innovation in newly public firms, *Journal of Financial Economics*, 111: 578–588.

Balsmeier, Benjamin, Lee Fleming, and Gustavo Manso, 2017. Independent boards and innovation. *Journal of Financial Economics*, 123: 536–557.

Benner, Mary J., and Michael Tushman, 2002. Process Management and Technological Innovation: A Longitudinal Study of the Photography and Paint Industries. Administrative Science Quarterly, 47: 676-706.

Bhaskarabhatla, Ajay, Luis Cabral, Deepak Hegde, Thomas Peeters, 2021. Are Inventors or Firms the Engines of Innovation?. *Management Science*, 67: 3899-3920.

Carnahan, Seth, and Deepak Somaya, 2013. Alumni effects and relational advantage: The impact on outsourcing when a buyer hires employees from a supplier’s competitors. *Academy of Management Journal*, 56: 1578–1600.

Cascio, Wayne F., 1991. Costing human resources: The financial impact of behavior in organizations. *Boston: PWS-Kent*.

Cassiman, Bruno, and Masako Ueda, 2006. Optimal project rejection and new firm start-ups. *Management Science*, 52: 262–275.

Chang, Xin, Kangkang Fu, Angie Low, and Wenrui Zhang, 2015. Non-executive employee stock options and corporate innovation. *Journal of Financial Economics*, 115: 168–188.

Chemmanur, Thomas J., Manish Gupta, and Karen Simonyan, 2020. Management Quality and Innovation in Private Firms and the IPO Market Rewards to Innovative Activity. *Entrepreneurship Theory and Practice*: forthcoming.

Chemmanur, Thomas J., Lei Kong, Karthik Krishnan, and Qianqian Yu, 2019. Top Management Human Capital, Inventor Mobility, and Corporate Innovation. *Journal of Financial and Quantitative Analysis*, 54: 2383-2422.

Chemmanur, Thomas, and Xuan Tian, 2017. Do anti-takeover provisions spur corporate innovation?. *Journal of Financial and Quantitative Analysis*, 53: 1163-1194.

Chen, Deqiu, Huasheng Gao, Yujing Ma, 2021. Human capital-driven acquisition: evidence from the inevitable disclosure doctrine. *Management Science* 67: 4643-4664.

Chu, Yongqiang, Xuan Tian, and Wenyu Wang, 2019. Corporate innovation along the supply chain. *Management Science*, 65: 2445-2466.

Coff, Russell W., 1999. When competitive advantage doesn't lead to performance: The resource-based view and stakeholder bargaining power. *Organization Science*, 10: 119-133.

Diamond, Douglas W., and Raghuram G. Rajan, 2000. A theory of bank capital. *Journal of Finance*, 55: 2431-2465.

Dokko, Gina, and Lori Rosenkopf, 2010. Social capital for hire? Mobility of technical professionals and firm influence in wireless standards committees. *Organization Science*, 21: 677–695.

Ederer, Florian, and Gustavo Manso, 2013. Is pay for performance detrimental to innovation?. *Management Science*, 59: 1496–1513.

Faleye, Olubunmi, Tunde Kovacs, and Anand Venkateswaran, 2014. Do better-connected CEOs innovate more?. *Journal of Financial and Quantitative Analysis*, 49: 1201–1225.

Eisfeldt, Andrea L., and Dimitris Papanikolaou, 2013. Organization capital and the cross-section of expected returns. *Journal of Finance*, 68: 1365-1406.

Fuller, Anne W., and Frank T. Rothaermel, 2012. When stars shine: The effects of faculty founders on new technology ventures. *Strategic Entrepreneurship Journal*, 6: 220–235.

Galasso, Alberto, and Timothy S. Simcoe, 2011. CEO overconfidence and innovation. *Management Science*, 57: 1469–1484.

Grant, Robert M., 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17: 109–122.

Groysberg, Boris, and Linda-Eling Lee, 2009. Hiring stars and their colleagues: Exploration and exploitation in professional service firms. *Organization Science*, 20: 740–758.

Hoisl, Karin, 2007. Tracing mobile inventors: The causality between inventor mobility and inventor productivity. *Research Policy*, 36: 619–636.

Hsu, Po-Hsuan, Yiqing Lü, Hong Wu, and Yuhai Xuan, 2024. Director job security and corporate innovation. *Journal of Financial and Quantitative Analysis*, 59: 652-689.

Hui, Kai Wai, Chuchu Liang, P. Eric Yeung, 2019. The effect of major customer concentration on firm profitability: competitive or collaborative?. *Review of Accounting Studies*, 24:189-229.

Irving, Paul J., Shawn Saeyeul Park, and Çelim Yildizha, 2016. Customer-base concentration, profitability, and the relationship life cycle. *The Accounting Review*, 91:883-906.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, 2018. Team-specific capital and innovation. *American Economic Review*, 108: 1034-1073.

Kacperczyk, Aleksandra J., 2013. Social influence and entrepreneurship: The effect of university peers on entrepreneurial entry. *Organization Science*, 24: 664–683.

Klasa, Sandy, Hernán Ortiz-Molina, Matthew Serfling, and Shweta Srinivasan, 2018. Protection of trade secrets and capital structure decisions. *Journal of Financial Economics*, 128: 266–286.

Kogut, Bruce, and Udo Zander, 1992. Knowledge of the firm, combinative capabilities, and replication of technology. *Organization Science*, 3: 383-397.

Kwon, Kiwook, and Deborah E. Rupp, 2013. High-performer turnover and firm performance: The moderating role of human capital investment and firm reputation. *Journal of Organizational Behavior*, 34: 129-150.

Liu, Kun, 2014. Human capital, social collaboration, and patent renewal within U.S. pharmaceutical firms. *Journal of Management*, 40: 616-636.

Liu, Tong, Yifei Mao, and Xuan Tian, 2017. The role of human capital: Evidence from patent generation. *Working paper*.

Mao, C., and C. Zhang, 2018. Managerial risk-taking incentive and firm innovation: Evidence from FAS 123R. *Journal of Financial and Quantitative Analysis*, 53: 867-898.

Manso, Gustavo, 2011. Motivating innovation. *Journal of Finance*, 66: 1823–1860.

Mayer, Roger C., Richard S. Warr, and Jing Zhao. Do pro-diversity policies improve corporate innovation?. *Financial Management*, 47: 617-650.

Nelson, Richard R., and Sidney G. Winter, 1982. An evolutionary theory of economic change. *Cambridge, MA: Harvard University Press*.

Paruchuri, Srikanth, 2010. Inter-organizational networks, intra-organizational networks, and impact of central inventors: A longitudinal study of pharmaceutical firms. *Organization Science*, 21: 63-80.

Patatoukas, Panos N., 2012. Customer-base concentration: Implications for firm performance and capital markets. *The Accounting Review*, 87: 363-392.

Raffiee, Joseph, 2017. Employee mobility and interfirm relationship transfer: Evidence from the mobility and client attachments of United States federal lobbyists, 1998–2014. *Strategic Management Journal*, 38: 2019–2040.

Shaw, Jason D., Nina Gupta, and John E. Delery, 2005. Alternative conceptualizations of the relationship between voluntary turnover and organizational performance. *Academy of Management Journal*, 48: 50-68.

Somaya, Deepak, Ian O. Williamson, and Natalia Lorinkova, 2008. Gone but not lost: The different performance impacts of employee mobility between cooperators versus competitors. *Academy of Management Journal*, 51: 936–953.

Sørensen, Jespeer B., and Amanda J. Sharkey, 2014. Entrepreneurship as a mobility process. *American Sociological Review*, 79: 328–349.

Steffy, Brian D., and Steven D. Maurer, 1988. Conceptualizing the economic effectiveness of human resource activities. *Academy of Management Review*, 13: 271-286.

Stoffman Noah, Michael Woeppel, and M. Deniz Yavuz, 2022. Small innovators: no risk, no reward. *Journal of Accounting and Economics*, 74: in press.

Sunder, Jayanthi, Shyam V. Sunder, and Jingjing Zhang, 2017. Pilot CEOs and corporate innovation. *Journal of Financial Economics*, 123: 209–224.

Tzabbar Daniel, and Rebecca R. Kehoe, 2014. Can opportunity emerge from disarray? An examination of exploration and exploitation following star scientist turnover. *Journal of Management*, 40: 449-482.

Wang, Qinyu Ryan, and Yanfeng Zheng, 2022. Nest without birds: Inventor mobility and the left-behind patents. *Research Policy*, 51: 104485.

Wright, Patrick M., Russell Coff, and Thomas P. Moliterno, 2014. Strategic human capital: Crossing the great divide. *Journal of Management*, 40(2), 353–370.

Table 1
Summary statistics

This table presents the summary statistics of the variables used in the baseline model (Equation (1)). The sample comprises all firm-year observations in Compustat/CRSP from 1980 to 2016 that meet the following criteria: 1) the firm is headquartered in the U.S.; 2) the firm has positive net sales and book equities, and the value of book assets exceeds 10 million dollars; 3) the firm is not in the finance (SIC code 6000–6999) or utilities (SIC code 4900–4999) industries; 4) the firm's stock is traded on the NYSE, NASDAQ, or AMEX, with a share code of either 10 or 11 in CRSP; and 5) the firm filed at least one patent over the five-year period up to year t. Panel A shows the explanatory variables in Equation (1), whereas Panels B and C display the dependent variables measuring innovation productivity and innovation style, respectively. Panel D shows that correlations between the explanatory variables listed in Panel A. There are 62,360 firm-year observations with non-missing values for the variables listed in Panel A. For citation-based measures of innovation productivity, the sample period ends in 2011 instead of 2016. All ratios are winsorized at the 1st and 99th percentiles. Detailed variable definitions are provided in Appendix B.

Panel A: Explanatory variables

Variables	*N*	*Mean*	*STD*	*Min*	*25%*	*Median*	*75%*	*Max*
Inventor concentration	62,360	0.252	0.285	0	0.048	0.147	0.333	1
# patents over past 5 years	62,360	151	928	1	2	8	37	39,374
# citations over past 5 years	62,360	289	1,916	0	2.750	12.400	65	122,403
Book assets (million $)	62,360	6,209	30,016	10	105	419	2,378	1,079,500
Tobin's Q	62,360	2.080	1.636	0.624	1.116	1.533	2.342	10.252
Profitability	62,360	0.067	0.196	-0.807	0.038	0.116	0.172	0.367
R&D intensity	62,360	0.073	0.109	0	0.003	0.031	0.096	0.617
% exploitative patents in past 5 years	62,360	0.150	0.206	0	0	0.067	0.238	1
% exploratory patents in past 5 years	62,360	0.688	0.298	0	0.487	0.734	1	1

Panel B: Measures of firms' innovation productivity

Variables	*N*	*Mean*	*STD*	*Min*	*25%*	*Median*	*75%*	*Max*
# patents in $t+1$	54,922	29.576	190.073	0	0	1	8	9,814
# patents in $t+2$	49,831	28.675	164.070	0	0	1	8	9,015
# patents in $t+3$	46,524	30.327	169.468	0	0	1	9	9,015
# citations in $t+1$	48,591	48.047	315.031	0	0	1.400	11.859	20,361
# citations in $t+2$	44,323	50.535	338.369	0	0	1.333	12.333	20,739
# citations in $t+3$	41,413	55.439	382.426	0	0	1.400	13.367	27,815

Panel C: Measures of firms' innovation style

Variables	*N*	*Mean*	*STD*	*Min*	*25%*	*Median*	*75%*	*Max*
% exploitative patents in *t+1*	54,922	0.145	0.262	0	0	0	0.200	1
% exploitative patents in *t+2*	49,831	0.149	0.264	0	0	0	0.200	1
% exploitative patents in *t+3*	46,524	0.153	0.265	0	0	0	0.217	1
% exploratory patents in *t+1*	54,922	0.352	0.392	0	0	0.200	0.667	1
% exploratory patents in *t+2*	49,831	0.343	0.387	0	0	0.175	0.667	1
% exploratory patents in *t+3*	46,524	0.337	0.383	0	0	0.167	0.667	1

Panel D: Pearson correlations between explanatory variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Inventor concentration								
(2) Logarithm of # patents over past 5 years	-0.638							
(3) Logarithm of # citations over past 5 years	-0.632	0.947						
(4) Book assets (million $)	-0.358	0.594	0.522					
(5) Tobin's Q	-0.046	0.018	0.088	-0.176				
(6) Profitability	-0.004	0.110	0.060	0.391	-0.204			
(7) R&D intensity	-0.129	0.047	0.105	-0.358	0.368	-0.664		
(8) % exploitative patents in past 5 years	-0.232	0.302	0.330	0.063	0.130	-0.168	0.205	
(9) % exploratory patents in past 5 years	0.350	-0.470	-0.496	-0.133	-0.126	0.132	-0.183	-0.789

Table 2
Inventor-base concentration and innovation productivity

This table examines how firms' inventor-base concentration impacts innovation output in subsequent years (Equation (1)). The sample includes the firm-year observations described in Table 1. In columns (1) – (3), the dependent variables are the raw number of patents filed in years *t+1*, *t+2*, and *t+3*, respectively. In columns (4) – (6), the dependent variables are the citation-weighted number of patents filed in years *t+1*, *t+2*, and *t+3*, respectively, where the weight is the number of citations received over the five-year period after the award date, scaled by the median number of citations received by patents filed in the same technological class-year cell. Detailed variable definitions are provided in Appendix B. All regressions control for firm and year fixed effects. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Logarithm of # patents			*Logarithm of # citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	0.424***	0.321***	0.282***	0.394***	0.289***	0.271***
	(0.028)	(0.032)	(0.036)	(0.038)	(0.042)	(0.046)
Logarithm of # patents in past 5 years	0.530***	0.414***	0.321***			
	(0.013)	(0.015)	(0.017)			
Logarithm of # citations in past 5 years				0.422***	0.336***	0.256***
				(0.014)	(0.016)	(0.017)
Logarithm of book assets	0.173***	0.179***	0.174***	0.250***	0.233***	0.214***
	(0.012)	(0.015)	(0.017)	(0.017)	(0.019)	(0.022)
Tobin's Q	0.056***	0.061***	0.058***	0.064***	0.067***	0.064***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)
Profitability	0.187***	0.212***	0.274***	0.187***	0.218***	0.331***
	(0.039)	(0.047)	(0.054)	(0.058)	(0.068)	(0.075)
R&D intensity	0.377***	0.326***	0.309**	0.701***	0.515***	0.459***
	(0.091)	(0.110)	(0.124)	(0.130)	(0.153)	(0.171)
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,922	49,831	46,524	48,591	44,323	41,413
Adjusted R^2	0.843	0.827	0.823	0.788	0.780	0.782

Table 3
Inventor-base concentration and firms' innovation style

This table examines the impact of inventor-base concentration on innovation style (exploitation versus exploration) in subsequent years. The sample includes the firm-year observations described in Table 1. In columns (1) – (3), the dependent variables are the percentage of explorative patents among all patents filed in years *t+1*, *t+2*, and *t+3*, respectively. In columns (4) – (6), the dependent variables are the percentage of exploratory patents among all patents filed in years *t+1*, *t+2*, and *t+3*, respectively. All regressions control for firm and year fixed effects. Detailed variable definitions are provided in Appendix B. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	% exploitative patents			*% exploratory patents*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	0.024***	0.021***	0.033***	-0.055***	-0.057***	-0.037**
	(0.007)	(0.008)	(0.008)	(0.015)	(0.015)	(0.015)
% exploitative patents in past 5 years	-0.176***	-0.172***	-0.153***	-0.011	0.009	0.017
	(0.015)	(0.015)	(0.017)	(0.019)	(0.020)	(0.020)
% exploratory patents in past 5 years	-0.080***	-0.053***	-0.021*	0.064***	0.038**	0.006
	(0.010)	(0.010)	(0.011)	(0.015)	(0.015)	(0.016)
Logarithm of # patents in past 5 years	0.070***	0.061***	0.053***	-0.042***	-0.046***	-0.045***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Logarithm of book assets	-0.001	0.005	0.008***	0.036***	0.026***	0.020***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
Tobin's Q	0.004***	0.007***	0.007***	0.009***	0.006***	0.003**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Profitability	0.028*	0.015	0.030*	0.057***	0.056***	0.065***
	(0.015)	(0.016)	(0.017)	(0.019)	(0.020)	(0.020)
R&D intensity	0.010	0.025	0.053	0.163***	0.117***	0.081*
	(0.032)	(0.034)	(0.037)	(0.039)	(0.041)	(0.042)
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,922	49,831	46,524	54,922	49,831	46,524
Adjusted R^2	0.342	0.352	0.359	0.252	0.258	0.267

Table 4
Inventor-base concentration and innovation productivity: Controlling for technology concentration and the scale of inventor base

This table replicates Table 2 with additional control variables, including *Technology concentration* (Panel A*), # of inventors in past five years* (Panel B), and both (Panel C). The purpose is to verify that inventor concentration is not merely a proxy for these variables. Firm characteristics in the baseline model are included but not reported. All regressions control for firm and year fixed effects. Detailed variable definitions are provided in Appendix B. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Technology concentration

	Logarithm of # Patents			*Logarithm of # Citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	0.288***	0.210***	0.190***	0.255***	0.178***	0.188***
	(0.027)	(0.031)	(0.035)	(0.038)	(0.042)	(0.046)
Technology concentration	0.368***	0.282***	0.218***	0.268***	0.227***	0.190***
	(0.028)	(0.035)	(0.040)	(0.041)	(0.047)	(0.051)
Other firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,920	49,829	46,523	48,589	44,321	41,412
Adjusted R^2	0.844	0.828	0.823	0.788	0.781	0.782

Panel B: The number of inventors in the firm

	Logarithm of # Patents			*Logarithm of # Citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	2.350**	2.984**	2.570*	3.251**	3.874*	3.274
	(0.936)	(1.222)	(1.332)	(1.570)	(2.032)	(2.142)
# of inventors in past	0.379***	0.326***	0.296***	0.448***	0.300***	0.218***
5 years	(0.024)	(0.028)	(0.032)	(0.025)	(0.030)	(0.033)
Other firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,922	49,831	46,524	48,591	44,323	41,413
Adjusted R^2	0.845	0.829	0.824	0.793	0.783	0.783

Panel C: Technology concentration and the number of inventors in the firm

	Logarithm of # Patents			*Logarithm of # Citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	0.682***	0.560***	0.519***	0.833***	0.560***	0.456***
	(0.040)	(0.048)	(0.055)	(0.048)	(0.057)	(0.064)
Technology concentration	0.321***	0.240***	0.179***	0.321***	0.258***	0.213***
	(0.028)	(0.034)	(0.039)	(0.039)	(0.047)	(0.051)
# of inventors in past	0.358***	0.310***	0.284***	0.458***	0.307***	0.224***

5 years	(0.023)	(0.028)	(0.032)	(0.025)	(0.030)	(0.033)
Other firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,920	49,829	46,523	48,589	44,321	41,412
Adjusted R^2	0.846	0.829	0.825	0.794	0.783	0.784

Table 5
Inventor-base concentration and innovation productivity: Excluding firms with low innovation production

This table replicates Table 2 using subsamples that exclude firms with low innovation production. The goal is to verify that the results in Table 2 are not solely driven by firms with a small scale of innovation production. Panel A uses a subsample that excludes firms where the number of patents filed in the past five-year period is lower than or equal to the 25th percentile among all firms in the same year. Panel B uses a subsample that excludes firms where the number of inventors who filed at least one patent in the past five-year period is lower than or equal to the 25th percentile among all firms in the same year. Firm characteristics from the baseline model are included but not reported. All regressions control for firm and year fixed effects. Detailed variable definitions are provided in Appendix B. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: subsample excluding firms with low patent output

	Logarithm of # patents			*Logarithm of # citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	0.432***	0.395***	0.427***	0.385***	0.372***	0.480***
	(0.069)	(0.083)	(0.091)	(0.094)	(0.106)	(0.116)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,465	36,309	34,127	34,771	32,120	30,218
Adjusted R^2	0.831	0.814	0.811	0.776	0.767	0.770

Panel A: subsample excluding firms with small inventor base

	Logarithm of # patents			*Logarithm of # citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	0.649***	0.636***	0.599***	0.549***	0.595***	0.716***
	(0.103)	(0.118)	(0.129)	(0.137)	(0.152)	(0.164)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	40,238	36,953	34,682	35,483	32,728	30,742
Adjusted R^2	0.832	0.815	0.812	0.777	0.769	0.771

Table 6
Inventor-base concentration and innovation productivity: Instrumental variable regressions

This table examines how firms' inventor-base concentration impacts innovation output in subsequent years, using an instrumental variable approach. The instrumental variable is a dummy variable indicating whether the firm's historical headquarters in year *t* was located in a state that adopted the Inevitable Disclosure Doctrine (IDD). The sample includes the firm-year observations described in Table 1. Panel A presents the results of the first-stage regressions, where the dependent variable is *Inventor concentration*. Panel B presents the results of the second-stage regressions. In columns (1) – (3) of Panel B, the dependent variables are the raw number of patents filed in years *t+1*, *t+2*, and *t+3*, respectively. In columns (4) – (6) of Panel B, the dependent variables are the citation-weighted number of patents filed in years *t+1*, *t+2*, and *t+3*, respectively, where the weight is the number of citations received over the five-year period after the award date, scaled by the median number of citations received by patents filed in the same technological class-year cell. Detailed variable definitions are provided in Appendix B. All regressions control for firm fixed effects. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: The first-stage regressions

	Inventor concentration	
	(1)	(2)
Dummy for headquarters in an IDD state	-0.021***	-0.015***
	(0.006)	(0.006)
Logarithm of # patents in past 5 years	-0.099***	
	(0.004)	
Logarithm of # citations in past 5 years		-0.082***
		(0.004)
Logarithm of book assets	0.0003	-0.010***
	(0.003)	(0.003)
Tobin's Q	0.004***	0.006***
	(0.001)	(0.001)
Profitability	0.017	0.015
	(0.012)	(0.012)
R&D intensity	-0.044*	-0.083***
	(0.026)	(0.029)
Firm dummies	Yes	Yes
Observations	62,360	54,448
Adjusted R^2	0.757	0.757

Panel B: The second-stage regressions

	Logarithm of # patents			*Logarithm of # citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Inventor concentration	2.712**	3.556**	3.083*	4.252*	5.442*	4.651
(instrumented)	(1.099)	(1.514)	(1.627)	(2.215)	(3.249)	(3.308)
Logarithm of # patents in	0.743***	0.722***	0.584***			
past 5 years	(0.110)	(0.153)	(0.164)			
Logarithm of # citations in				0.731***	0.753***	0.607**
past 5 years				(0.184)	(0.272)	(0.275)

Logarithm of book assets	0.102***	0.106***	0.101***	0.204***	0.194***	0.166***
	(0.012)	(0.016)	(0.018)	(0.030)	(0.038)	(0.036)
Tobin's Q	0.043***	0.047***	0.045***	0.040***	0.039*	0.039*
	(0.005)	(0.007)	(0.008)	(0.014)	(0.020)	(0.021)
Profitability	0.237***	0.268***	0.345***	0.215***	0.261***	0.388***
	(0.048)	(0.060)	(0.063)	(0.076)	(0.093)	(0.094)
R&D intensity	0.461***	0.486***	0.465***	1.003***	0.998**	0.884**
	(0.121)	(0.161)	(0.173)	(0.260)	(0.393)	(0.404)
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,922	49,831	46,524	48,591	44,323	41,413
Adjusted R^2	0.796	0.737	0.756	0.685	0.603	0.658

Table 7
Inventor-base concentration and inventors' human capital quality

This table examines the association between inventor-base concentration and inventors' human capital quality (Equation (2)). The sample includes inventor-year observations. For each observation in the firm-year sample described in Table 1, we identify inventors who filed at least one patent over the five-year period up to year t. In column (1), the dependent variable is the raw number of patents filed by the inventor over their entire career up to year *t*. In column (2), the dependent variable is the citation-weighted number of patents filed by the inventor over their entire career up to year *t*, where the weight is the number of citations received over the five-year period after the award date, scaled by the median number of citations received by patents filed in the same technological class-year cell. In column (3), the dependent variable is a dummy variable indicating whether the inventor is a star in year *t*. A star inventor in year *t* is among the top 5% of inventors in terms of the citation-weighted number of patents filed over their entire career up to year *t*. Detailed variable definitions are provided in Appendix B. All regressions control for inventor and year fixed effects. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	# patents per inventor	*# citations per inventor*	*star*
	(1)	(2)	(3)
Firm characteristics:			
Inventor concentration	0.128***	-0.088**	0.005
	(0.028)	(0.042)	(0.011)
Logarithm of # patents in past 5 years	0.041***	0.004	0.006***
	(0.005)	(0.005)	(0.001)
Logarithm of book assets	-0.026***	-0.012**	-0.006***
	(0.006)	(0.005)	(0.001)
Tobin's Q	-0.005**	0.015***	0.003***
	(0.002)	(0.003)	(0.001)
Profitability	-0.0001	-0.100**	-0.022***
	(0.019)	(0.042)	(0.008)
R&D intensity	-0.042	-0.187***	-0.055***
	(0.042)	(0.066)	(0.015)
Inventor characteristics:			
Inventor seniority	0.187***	0.214***	0.058***
	(0.004)	(0.010)	(0.002)
Inventor dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Observations	6,562,482	5,071,934	5,071,934
Adjusted R^2	0.912	0.710	0.608

Table 8
Inventor-base concentration and inventor retention

This table examines the impact of inventor-base concentration on inventor retention in subsequent years (Equation (3)). The sample includes inventor-year observations. For each observation in the firm-year sample described in Table 1, we identify inventors who filed at least one patent over the five-year period up to year *t*. The dependent variables are dummy variables indicating whether the inventor remains with the firm in years *t+1*, *t+2*, and *t+3*, respectively. Detailed variable definitions are provided in Appendix B. All regressions control for inventor and year fixed effects. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Inventor stay		
	t+1	*t+2*	*t+3*
	(1)	(2)	(3)
Firm characteristics:			
Inventor concentration	0.157***	0.174***	0.182***
	(0.036)	(0.041)	(0.044)
Logarithm of # patents in past 5 years	0.041***	0.041***	0.039***
	(0.005)	(0.005)	(0.006)
Logarithm of book assets	-0.017***	-0.015***	-0.011**
	(0.004)	(0.004)	(0.005)
Tobin's Q	0.010***	0.014***	0.017***
	(0.002)	(0.002)	(0.002)
Profitability	0.020	0.039	0.050
	(0.023)	(0.027)	(0.031)
R&D intensity	-0.062	-0.105*	-0.109*
	(0.050)	(0.059)	(0.062)
Inventor characteristics:			
Logarithm of # patents by inventor	-0.036***	-0.012***	0.006**
	(0.003)	(0.002)	(0.003)
Logarithm of inventor tenure	-0.037***	-0.057***	-0.077***
	(0.003)	(0.003)	(0.004)
Inventor dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Observations	6,082,361	5,655,607	5,189,345
Adjusted R^2	0.352	0.383	0.415

Table 9
Inventor-base concentration and inventor-team retention

This table examines the impact of inventor-base concentration on the retention of a focal inventor's significant collaborator in subsequent years (Equation (4)). The sample includes inventor pair-year observations. For each observation in the firm-year sample described in Table 1, we identify focal inventors who filed at least one patent over the five-year period up to year *t*. Another inventor is deemed the focal inventor's significant collaborator if they are the most frequent collaborator of the focal inventor over the past five years. The dependent variables are dummy variables indicating whether the significant collaborator remains with the firm in years *t+1*, *t+2*, and *t+3*, respectively. Detailed variable definitions are provided in Appendix B. All regressions control for firm and year fixed effects. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Significant collaborator stay		
	t+1	*t+2*	*t+3*
	(1)	(2)	(3)
Firm characteristics:			
Inventor concentration	0.066*	0.105***	0.082***
	(0.037)	(0.036)	(0.030)
Logarithm of # patents in past 5 years	0.025***	0.027***	0.021***
	(0.006)	(0.005)	(0.004)
Logarithm of book assets	0.002	0.003	-0.002
	(0.004)	(0.003)	(0.003)
Tobin's Q	-0.0001	0.002**	0.003***
	(0.001)	(0.001)	(0.001)
Profitability	-0.002	0.008	0.025*
	(0.010)	(0.015)	(0.014)
R&D intensity	0.037*	-0.010	-0.003
	(0.021)	(0.028)	(0.024)
Inventor characteristics:			
Logarithm of # patents by significant	-0.017***	-0.007**	-0.003
collaborator	(0.004)	(0.003)	(0.003)
Logarithm of significant collaborator tenure	0.001	-0.003	-0.007***
	(0.002)	(0.002)	(0.002)
Logarithm of # patents by inventor	-0.012***	-0.015***	-0.012***
	(0.002)	(0.002)	(0.002)
Logarithm of inventor tenure	-0.013***	-0.013***	-0.016***
	(0.002)	(0.002)	(0.002)
Inventor-pair characteristics:			
Logarithm of frequency of collaboration	0.035***	0.022***	0.015***
	(0.003)	(0.002)	(0.002)
Logarithm of inventor-collaborator distance	-0.001***	-0.001***	-0.001***
	(0.0003)	(0.0002)	(0.0003)
Inventor-pair dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes

Observations	10,568,126	10,322,032	10,008,782
Adjusted R^2	0.600	0.660	0.718

Table 10
Alternative measure of inventor-base concentration

This table presents the results of the robustness tests using an alternative measure of inventor-base concentration. The alternative measure is calculated as the proportion of patents filed by the most prolific inventor within the firm over the past five years relative to the total patents filed by the firm over the same period. Panels A and B replicate Tables 2 and 3, respectively. Firm characteristics are included but not reported. All regressions control for firm and year fixed effects. Detailed variable definitions are provided in Appendix B. The standard errors of the estimated coefficients (in parentheses) are heteroscedasticity-robust and allow for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Firms' innovation productivity

	Logarithm of # Patents			*Logarithm of # Citations*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Top inventor's share	0.288***	0.210***	0.190***	0.255***	0.178***	0.188***
	(0.027)	(0.031)	(0.035)	(0.038)	(0.042)	(0.046)
Other firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,922	49,831	46,524	48,591	44,323	41,413
Adjusted R^2	0.842	0.826	0.823	0.787	0.780	0.782

Panel B: Firms' innovation style (exploitation versus exploration)

	% exploitative patents			*% exploratory patents*		
	t+1	*t+2*	*t+3*	*t+1*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)	(5)	(6)
Top inventor's share	0.027***	0.019**	0.029***	-0.057***	-0.056***	-0.033**
	(0.008)	(0.008)	(0.009)	(0.015)	(0.016)	(0.016)
Other firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,922	49,831	46,524	54,922	49,831	46,524
Adjusted R^2	0.342	0.352	0.359	0.252	0.258	0.267