Measuring the Characteristics and Employment Dynamics of U.S. Inventors*

Ufuk Akcigit† Nathan Goldschlag‡

March 25, 2023

Abstract

Innovation is a key driver of long run economic growth. Studying innovation requires a clear view of the characteristics and behavior of the individuals that create new ideas. A general lack of rich, large-scale data has constrained such analyses. We address this by introducing a new dataset linking patent inventors to survey, census, and administrative microdata at the U.S. Census Bureau. We use this data to provide a first look at the demographic characteristics, employer characteristics, earnings, and employment dynamics of inventors. These linkages, which will be available to researchers with approved access, dramatically increases the scope of what can be learned about inventors and innovative activity.

JEL Codes: O3, O4
Keywords: Inventors, innovation, R&D, firms, dynamism, reallocation

*Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. DRB Approval Number(s): CBDRB-FY20-CES007-004, CBDRB-FY21-CES007-004, CBDRB-FY22-CES008-008, CBRDB-FY23-CES020-001, CBRDB-FY23-CES020-002. DMS Project Number 7083300. We thank Joseph Staudt, Shawn Klimek, Emek Basker, participants at the 2022 Research Data Center Annual Conference and USPTO Chief Economist Seminar for helpful comments and suggestions.

†University of Chicago, NBER, CEPR

‡U.S. Census Bureau

1 Introduction

A large literature characterizes innovation using measures of the R&D and patenting activity of firms (Griliches, 1992, 1998; Hall, 1996). While inventors themselves have long been of interest (Winston, 1937; Schmookler, 1957; Lamoreaux and Sokoloff, 1999), only recently have researchers been able to assemble large administrative datasets capable of characterizing the population of inventors. Recent work using such datasets has shed light on inventor's contribution to long run growth and the relationship between innovation and income inequality (Akcigit et al., 2017b), the relationship between talent, human capital, and innovation (Akcigit et al., 2020), the individual returns to innovative activity (Toivanen and Väänänen, 2012; Aghion et al., 2018; Kline et al., 2019), the role of team-specific human capital (Jaravel et al., 2018), and the processes by which individuals become inventors (Aghion et al., 2017; Celik, 2023), the role of parental resources, contributions of IQ, and the role of childhood environments (Bell et al., 2019; Aghion et al., 2023). In this paper we extend the reach of such analyses by introducing a dataset linking U.S. inventors associated with patents granted since 2000 to Census Bureau data on individuals and their employment histories. We use that data to provide a first look at the demographic characteristics, employer characteristics, earnings, and employment dynamics of U.S. inventors.

Linking individual inventors to the Census Bureau's extensive survey and administrative data sources will allow researchers to unpack the "black box" of innovative activity. Historically, one of the best measured inputs to innovative activity is R&D dollars spent by large firms. This type of information was collected via survey instruments such as the Business Research & Development and Innovation Survey (BRDIS) and its predecessor the Survey of Industrial Research & Development (SIRD) (Foster et al., 2020). The utilization of human capital inputs to the innovation process is less well understood. Linking patent inventors to survey, census, and administrative employee-employer information dramatically expands the types of analyses one can pursue regarding the role of individuals in the inventive process. These data, which we develop and describe in this paper, will be available to researchers with approved access via the Federal Statistical Research Data Centers.

Creating quality links between inventors and Census information on individuals is complicated by the limited amount of personally identifiable information available within patent documents. Patent documents contain the inventor's name and geographic location (city and state). Inventors with common names residing in large cities will be difficult to match uniquely. To overcome this, we utilize a triangulation methodology that simultaneously considers inventor-individual matches, assignee-firm matches, and individual-firm employment relationships (Graham et al., 2018; Dreisigmeyer et al., 2018). Using this methodology, we are able to collect the demographic characteristics and employment histories for over 760 thousand inventors on U.S. patents granted between 2000 and 2016.

We utilize these data to present a series of stylized facts about inventors. We group our empirical findings into those dealing with (i) demographic characteristics, (ii) employer characteristics, (iii) earnings, and (iv) employment dynamics. Each set of findings speaks to a distinct set of policy relevant questions in the literature. The first group, focused on demographics, provides measures of the shifting demographic composition of the inventor population over time by characteristics such as age, sex, race, ethnicity, and foreign born status. These patterns are key to evaluating policies aimed at increasing the representation of specific groups. In our data we find familiar patterns of under representation of women and minorities among inventors (Akcigit et al., 2017a; Bell et al., 2019; USPTO, 2019; Miranda and Zolas, 2017; Cook and Kongcharoen, 2010). We also find significant changes in these patterns over time—the share of Asian inventors rose by over two thirds from 2000 to 2016. Finally, our analysis of inventors born outside of the U.S. speaks to how immigration policies might affect aggregate innovative activity.

Our second group of findings concern the characteristics of firms that employ inventors. These firms are especially important for aggregate innovative outcomes because they choose how inventive talent is utilized (Akcigit et al., 2020; Akcigit and Goldschlag, 2023). We find that inventor productivity varies significantly between different types of firms. Inventors are increasingly concentrated in old, large incumbent firms while inventors employed by young firms tend to produce more impactful innovations. A better understanding of firms that employ inventors not only provides a more complete picture of the set of firms engaged in innovative activity, but it also allows for the design of more targeted innovation policies.

The third set of findings characterize inventor earnings. One way that inventors capture the returns of their inventive output is through earnings received from their employer. The distribution of earnings among inventors may have important implications for individual incentives to innovate and how returns are allocated between firms and inventors (Toivanen and Väänänen, 2012; Aghion et al., 2018; Kline et al., 2019). Moreover, a clear view of where inventors sit on the earnings distribution may have implications for how tax policies affect innovative activity. We find that the inventor earnings distribution is very skewed. Over 88 percent of super star inventors, those with the most impactful patents, are in the top decile of the earnings distribution and nearly 19 percent are in the top one percent of the earnings distribution.

Finally, our last set of empirical results describe the employment dynamics of inventors. A large literature has documented a secular decline in business dynamism and employment flows (Decker et al., 2014; Hyatt and Spletzer, 2013). These patterns have important implications for the efficient allocation of resources in the economy. The reallocation of inventors may be especially important if knowledge creation and innovative activity have spillover effects. We find that over time inventors are less likely to start a firm and less likely to change jobs. We also find that inventors are becoming increasingly concentrated geographically, which has implications for place-based innovation policies.

The following summarizes the basic empirical patterns in the data.

Inventor Demographics

1. **Females are under represented among inventors, especially on a citation weighted basis.** Females account for less than 12% of inventors, a share that is rising and tends to be higher among young inventors. The share of citations accounted for by female inventors lags behind their share of the inventor population. (See Figure 1)

2. **Many inventors are foreign born.** Over 30% of inventors are foreign born and we see a rise and fall of foreign born share among young inventors. China and India account for an increasing share of foreign born inventors, rising from 25% to 40% by 2016. (See Figure 2, 3)

3. **Inventors are getting older.** The average age of inventors rises from 43 to 46 over this period. The share of young inventors falls through 2011, but begins to rise thereafter. (See Figure 4)

4. **African Americans are significantly under represented among inventors, Asians inventors are increasingly common.** We find under representation of most minorities, especially for African Americans, who account for less than two percent of inventors. Asians, on the other hand, rose from 13% to 22% of inventors. (See Table 4)

5. **Representation of different demographic groups varies significantly by sector.** Female inventors have greater representation in the Health Care and Social Assistance and Education sectors. Foreign born inventors are more common in the Information, Education, and Professional/Scientific Services sectors. The share of young inventors is highest in the Information sector. (See Table 5)

Employer Characteristics

6. **Inventors work at older, larger firms.** Over 68% of inventors work in firms over 20 years old. Almost 63% of inventors work in large firms with at least 1000 employees. Akcigit and Goldschlag (2023) show that the share of inventors at incumbent firms rose from 49% to 58% between 2000 and 2016. This pattern appears to go back to the 1970s. The share of inventors on patent grants with the largest assignees between 1980 and 2018 rose from about 34% to 47% percent. (See Table 6, 7 and Figure 7)

7. **Inventors, especially the most productive, are much less likely to work at young firms.** The share of inventors working at firms $\leq$ 5 years old fell by almost half over this period from 15% to under 8%. This share fell the most among super star inventors, those with the most impactful patents. (See Figure 6, 8)

8. **Inventors at young firms produce more impactful patents.** Inventors at young, medium size firms tend to have the highest impact patents. Inventors at older, smaller firms tend to have the lowest impact patents. (Table 8)

Inventor Earnings

9. **Inventors earnings is highly skewed, especially for super star inventors.** About 63% of all inventors, and 88% of super star inventors, are in the top 10% of the earnings distribution. Almost 8% of all and 19% of super star inventors are in the top 1%. (See Table 9)

10. **Inventor earnings is closely tied to inventive productivity.** Inventors in the top 10% of the inventor earnings distribution have average citations per grant about 0.5 log points higher than inventors in the bottom 10%. (See Figure 10)

Inventor Employment Dynamics

11. **Inventors are less likely to switch jobs over time.** The hire and separation rates for inventors fell from about 6% and 7% respectively in 2000 to 4% in 2016. (See Figure 11)

12. **Inventors, especially super star inventors, are less likely to start a firm over time.** The probability a super star inventor becomes an entrepreneur fell by 57%. (See Figure 12)

13. **Inventors are increasingly geographically concentrated and less likely to change employment across state lines.** The share of inventors working in the 20 largest

counties by inventor count rose from 39% to over 47% between 2000 and 2016. The share of inventors switching employment across state lines fell from a peak of 4.6% in 2006 to 2.6% in 2016. (Figure 13, 14)

The paper proceeds as follows. Section 2 describes the matching procedures, quality of the resulting links, and the data used to describe the basic facts of U.S. inventors. Section 3 describes presents in greater detail the demographic characteristics, employer characteristics, earnings, and employment dynamics of inventors. Section 4 concludes.

2 Matching Inventors

The Census Bureau assigns disambiguated, anonymized, person-level identifiers called Protected Identification Keys (PIKs) to survey and administrative data that includes individuals. Matching patent inventors to PIKs allows us to link inventors to other Census Bureau microdata such as the Decennial Census, American Community Survey (ACS), the Longitudinal Employer Household Dynamics (LEHD), and W2 records. This paper introduces an inventor record-to-PIK database covering US-based inventors associated with patents granted between 2000 and 2016.[1] Patent inventor records are unique patent grant-inventor combinations found in the United States Patent and Trademark Office (USPTO) data. Inventors on patent grants that have more than one inventor are differentiated by a unique counter, the inventor sequence number, which captures the order an inventor appears in the patent document.

The bulk of our inventor record-PIK database is made up of matches drawn from the Business Dynamics Statistics of Patenting Firms (BDSPF) triangulation matching methodology.[2] Inventor record-PIK matches are generated by the BDSPF triangulation method when a "fully traingulated" match is made. This triangulation process helps resolve ambiguity in the person matches and firm matches, which often contain multiple false positives. A fully triangulated match occurs when, for a given patent, an inventor record-PIK match and an assignee-firm match are corroborated by an PIK-firm employment spell in the LEHD data. These represent the highest quality assignee-firm and inventor record-PIK matches in our data because matches for the inventor and assignee are validated by an employee-employer linkage in the LEHD data. With these matches we are able to assign PIKs to inventor records.

[1]The set of years available for the match is limited by the availability of comprehensive LEHD data. Foreign inventors are not at risk of matching Census Bureau data.

[2]See Graham et al. (2018) and Dreisigmeyer et al. (2018) for details on the patent triangulation match methodology and implementation.

We also incorporate unique matches output by the Person Validation System (PVS). The Census Bureau's PVS is used to assign an anonymous, person-level identifier to individuals.[3] Typically, the PVS matches an input file with personally identifiable information (PII) such as name, date of birth, and in some cases SSN, to a confidential reference file. The reference file maps individual PII to a unique PIK. One input of the BDSPF triangulation match are PVS inventor matches that rely on the relatively limited information available within patent documents—the person name, city, and state. The limited PII available for inventors produces relatively ambiguous matches. Many inventor records match to multiple PIKs. Numerous false positives from the PVS match is a feature, not a bug, for the triangulation match because additional assignee-firm and person-firm (job) links are used to disambiguate the relatively messy matches. In some cases, however, either when the geography is relatively small or the inventor name is relatively unique, the PVS will assign a unique PIK to an inventor record. When available, we incorporate these unique matches into our inventor record-PIK database.

Finally, we augment our inventor matches using disambiguated inventor identifiers from the PatentsView (PV) database. The PV database contains inventor identifiers intended to link inventor records across patent grants that represent the same individual (Monath et al., 2021). Those identifiers are constructed using inventor information such as name and geography along with co-invention networks. Inventor records with the same (or similar) name that appear across different patent documents with the same (or similar) team of inventors will be linked together. Importantly, our triangulated and PVS matches described above assign PIKs to inventor records, not disambiguaged groupings of inventor records that belong to the same, disambiguated person. Given that, we expect there to be inventor records associated with the same person, but that we were unable to assign a PIK (perhaps because we were unable to match the assignee to a firm in the Business Register). In these cases we can utilize the PV disambiguation to extend the reach of our inventor record-PIK matches. Whenever a PIK uniquely matches to a PV inventor identifier (InvID), and that InvID links to inventor records we were unable to match, we assign the PIK to those additional inventor records.[4]

Ultimately, of the 4.9 million US-based inventor records at risk of a match we are able to assign a PIK to 4.5 million (93 percent). Table 1 shows the distribution of match types. The vast majority (67 percent) of our inventor record matches are derived from

[3]See Wagner and Lane (2014) for details.

[4]Note that we do not match additional PIKs using the PV information. The PV identifiers simply allow us to cover additional inventor records with our already matched set of PIKs. Note also that we only utilize PV identifiers to which our PIK matches map uniquely. That is, when an InvID links to multiple PIKs we do not use it in our matching.

BDSPF triangulation matches. About a quarter of our inventor record matches come from unique PVS matches and another 8 percent from extensions using the PV inventor identifiers.

There are some systematic differences in the characteristics of patents for which we are able to match an inventor record and patents we do not. Table 2 shows several patent-level metrics associated with matched and unmatched inventor records. On average, matched inventor records are associated with patents with more than twice as many citations made and nearly twice as many citations received. Matched inventor records are also associated with more independent claims and slightly larger teams (number of inventors). Differences in team size are likely driven by the fact that patents with more inventors are more likely to receive a match for at least one inventor record.

Whenever our BDSPF triangulation and PVS matching assigns a single PIK to multiple inventor records we necessarily disambiguate inventors in the patent data. This provides us with the opportunity to compare two disambuguated identifiers: the PIKs assigned by BDSPF triangualtion and PVS matching and the PV InvIDs. These two disambiguation methods are quite different but yield very similar groupings of inventor records. Since the groupings are not identical, there is a many-to-many mapping of InvIDs to PIKs. We are able to assess the ambiguity of the groupings in "both directions". That is, we can measure how often InvIDs map to multiple PIKs and how often PIKs map to multiple InvIDs.

The left panel of Table 3 shows the share of InvIDs that map to 1, 2, 3, 4, and 5 or more PIKs. As shown in the second column, about 94 percent of InvIDs map to a single PIK and almost 5 percent map to two PIKs. Less than 1 percent of InvIDs map to more than two PIKs. Among the InvIDs that map to multiple PIKs, we can evaluate how concentrated that mapping is by measuring how many inventor records are associated with the modal InvID-PIK pair. For example, if an InvID maps to two PIKs, one of those InvID-PIK pairs could account of over 90 percent of inventor records for that InvID or they could be split roughly equally. In the third column of Table 3 we see that among the InvIDs that map to two PIKs, one of those PIKs accounts for about 71 percent of the inventor records. Even among the InvIDs that map to five or more PIKs the inventor records tend to be concentrated in one of the PIKs—the PIK with the most inventor records accounts for over half of inventor records.

The right panel of Table 3 describes ambiguity in the "other direction". We find slightly less ambiguity in the mapping of PIKs to InvIDs. Nearly 97 percent of PIKs map to a single InvID and about 2.7 percent map to two InvIDs. Again, fewer than one percent of PIKs map to more than two InvIDs. We see a similar pattern of concentration in the inventor records. The InvID match with the most inventor records accounts for 72 percent of records for PIKs that map to two InvIDs and about 46 percent among PIKs that map to 5 or more InvIDs.

Our inventor record-PIK database provides a high quality bridge between inventor records in the patent data to unique person identifiers in the Census Bureau data. With those links we are able to combine to survey, census, and administrative data on individual inventors. The overwhelming majority of our PIK matches come from high quality BDSPF triangulation matches. Moreover, we observe substantial overlap between the PatentsView InvIDs and our PIK matches, a positive signal about both the quality of the PatentsView disambiguation algorithm and the PIK matches. The inventor record-PIK database will be made available to researchers with approved access.

2.1 Inventor Jobs Panel

In order to measure the demographics and employment dynamics of inventors we link our inventor-PIK matches to the LEHD data to create an inventor jobs panel. This panel, which we term the Inventor Employment History panel, contains the full employment history of all persons linked to at least one patent grant in our sample period. The LEHD data is derived from quarterly state unemployment insurance (UI) records covering about 95 percent of private sector employment as well as state and local government employment. Not all states are available in the LEHD data every period. We restrict to 45 states with employment data starting in 2004 or earlier. Our panel includes only dominant, beginning-of-quarter (BoQ) jobs from 2000Q1 to 2016Q3.[5] After 2000, the addition of states to the frame introduces only minimal bias in national of job-to-job statistics (Henderson and Hyatt, 2012) and have very little impact on measures of inequality (Abowd et al., 2017).

To each quarterly job observation we attach measures of inventive activity by combining our inventor record-PIK links with patent information from PatentsView. Though our jobs panel ends in 2016, we use patent grants through 2019, extending the range of our citations received measures. Each quarter we track the number of patent and citation weighted applications and grants associated with the inventor. Our preferred patent citation measure counts external citations received, or citations associated with a different assignee, in the five years after a patent is granted. Separately, in order to identify more incremental innovations, we track the number of self citations received in the same five

[5] A beginning-of-quarter job is defined as a job for which the individual has positive earnings in both t and $t-1$. A dominant beginning-of-quarter job is the highest earning beginning-of-quarter job for the individual within a given quarter.

year window. Finally, we use the independent claim counts developed by Marco et al. (2019) to measure the scope of each patent.

3 Results

3.1 Inventor Demographics

Consistent with previous work, we find under representation of females among inventors. Figure 1 shows the female share of grant active inventors between 2000 and 2016.[6] The share of female inventors rose from about 8 percent in 2000 to over 11 percent in 2016. This is less than the 13.1 percent found by Bell et al. (2019), similar to the 10 percent reported by USPTO (2019), and higher than the 7.9 percent found by Miranda and Zolas (2017).[7] Among young inventors, those less than 36 years old, the female share rose from about 11 percent to over 16 percent—still far below parity. We also find that female inventors tend to be associated with lower-citation patents. Females account for about 10 percent of inventors but only about 7 percent of citations received.[8]

Our data also allow us to identify whether an inventor was born outside of the United States. Figure 2 shows the share of grant active inventors each year that are foreign born, which rises from about 24 percent to almost 35 percent. In contrast to female inventors, foreign born inventors tend to be on higher-citation patents. The foreign born share of citations received rises from about 30 percent in 2000 to nearly 40 percent in 2013. We also find a decline in share of young inventors that are foreign born, initially rising from 25 percent to over 30 percent, but then declining back to 25 percent after 2009.

In addition to foreign born status, we also observe the origin region, and in some cases country, of foreign born inventors. In Figure 3 we show how the share of foreign born inventors by origin region changed from 2000 and 2016. We see that the share of foreign born inventors from China and India rose significantly over this period. Combined, China and India went from just over 25 percent of all foreign born inventors in 2000 to over 40 percent in 2016. Despite the count of inventors from other Asian countries

[6]Grant active inventors are those that receive at least one patent grant in the period.

[7]Miranda and Zolas (2017), who focus on the role of household and non-employer inventor activity, utilize PVS matched inventors without the benefit of the Graham et al. (2018) triangulation match. Moreover, the demographic shares reported by Miranda and Zolas (2017) are also not directly comparable to those reported here because they reflect patent-level averaged characteristics and not percentages of inventors.

[8]The citation series is censored at 2013 because we require five years of subsequent citation data. Lower shares of citations, relative to share of inventors, could be related to technology field and industry composition.

and Europe rising over this period, both fall as a share due to the significant increase in flows from China and India.

The age of the U.S. labor force has risen significantly over the last few decades. The share of workers age 55 or older rose from 13.1 percent in 2000 to 23.6 percent in 2020.[9] Inventors show similar patterns of aging over time. Figure 4 shows the average age of inventors rose by about 3 years between 2000 and 2016. A complementary measure, the share of inventors less than 36 years old, fell by about 3 points over the same period. Among the entire civilian labor force, the share of workers less than 35 years old fell from 38.8 to 35.3 between 2000 and 2020. The share of young inventors after 2010 appears to be rising as the average age plateaus.

Our data provide a unique opportunity to summarize the self-reported race and ethnicity of inventors. Previous work aimed at describing the race of inventors relied on surname-based algorithms or matching curated lists of inventors (Cook and Kongcharoen, 2010; Cao et al., 2021; Breschi et al., 2017). We find substantial changes to the composition of inventors over this period. Figure 5 summarizes these changes by showing, for each race and ethnicity group, the ratio of the share of the inventor population in 2000 and 2016 to the group's share of the 2010 population.[10] In both years the Black, American Indian or Alaska Native Alone (AIAN), Hispanics, and individuals with two or more race categories are under represented relative to their population share. The Black or African American Alone group, for instance, accounted for about 12.6 percent of the population in 2010 and only about 1.2 percent of inventors in 2000 (a ratio of 0.1 in the figure). Asians, on the other hand, had an inventor share in 2000 about 2.6 times their population share. This rose to over 4.6 times their 2010 population share by 2016.

In terms of a relative propensities, Figure 5 shows that the Native Hawaiian or Other Pacific Islander Alone (NHOPI) became more likely to be inventors from 2000 to 2016. To aid intuition about the racial and ethnic composition of inventors, we also consider the overall distribution of inventors across groups. Table 4 shows the inventor shares for each group in each period. We see that NHOPI accounted for about 0.06 percent of inventors in 2000 and rose to 0.17 percent by 2016. The increased representation of Asians is driven by a 9.3 percentage point increase (from 12.6 to 21.9). The share of the White Alone group fell by a similar amount, 10.7 percentage points (from 85.3 to 74.6). By ethnicity, the share of Hispanic inventors also rose from 2 to 3.7 percent, but remains

[9] See BLS Employment Projections, `https://www.bls.gov/emp/tables/civilian-labor-force-summary.htm` (accessed 09/2022).

[10] See `https://www.census.gov/content/dam/Census/library/publications/2011/dec/c2010br-02.pdf` for 2010 decennial race and ethnicity population shares.

well below the 2010 population share of 16.4 percent. The patterns we observe in the racial composition of inventors are consistent with previous studies showing the under representation of minority groups among inventors (Cook and Kongcharoen, 2010; Bell et al., 2019).[11]

Having the employment histories of inventors allows us to simultaneously describe worker and employer characteristics. The demographic composition of inventors varies significantly by sector. Table 5 shows the share of female, foreign born, young ($\leq$ 35), older ($\geq$ 56), and non-White inventors for a subset of sectors. The highest share of female inventors is in Health Care and Social Assistance (17.8%) and Educational Services (14.8%) and Finance, Insurance, Real Estate, Rental, and Leasing (14.3%). The inventor female share falls to 5 percent in Mining, Utilities, and Construction. The share of foreign born inventors is over 40 percent in Information and Educational Services. Information is also the sector with the highest share of young inventors (23%). Finally, non-White inventors are most common in Information (29.6%) and least likely to be in Mining, Utilities, and Construction (15.2%). These patterns demonstrate the importance of considering sectoral composition when describing inventor demographic characteristics.

3.2 Employer Characteristics

Just as the composition of R&D expenditures across firms of different types has important implications for aggregate innovative activity, so too does the distribution of inventors (Akcigit and Goldschlag, 2023). Table 6 shows the distribution of inventor jobs by firm age and inventor age. Roughly 10 percent of inventor jobs are in young firms ($\leq$ 5 years old). The majority of inventor jobs (68%) are in firms more than 20 years old. Consistent with the literate showing that young workers are over represented among young firms (Ouimet and Zarutskie, 2014), the share of inventor jobs at young firms is highest among young workers, falling by nearly half (15% to 8%) between those 25 or younger and those 56 or older. As shown in Figure 6, over our sample period the share of inventors at young firms declined from roughly 15 percent in the early 2000s to less than 8 percent by 2016.[12] Akcigit and Goldschlag (2023) show that the share of inventors in old, larger incumbent firms has risen substantially from 2000 to 2016.

[11] Bell et al. (2019) do not observe race and ethnicity in their tax data sample, but report inventor rates by race and ethnicity among the NYC school sample.

[12] This pattern is not be unique to inventors and can be seen for all workers in the industries in which inventors are most frequent. For example, using a 1% sample of LEHD workers weighted by the industry composition of inventors, we find the share of workers at young firm that falls from roughly 14% in 2000 to just over 8% in 2016.

This pattern of increasing concentration of inventors among incumbent firms appears to go back at least to the 1970s. Figure 7 shows the share of inventors in the PatentsView data on patent grants with an assignee that is among the top one percent of assignees by inventor count.[13] We find that the share of inventors on patent grants with the largest assignees between 1980 and 2018 rose by almost 13 percentage points from about 34 percent to 47 percent. Over that period, the count of unique inventors an assignee must be associated with to be classified as the top one percent rose from 81 to 154. These patterns suggest that the rising share of inventors employed by incumbents is a phenomenon that predates our sample period.

Compared to firm age, the distribution of inventors across firm size groups does not exhibit significant differences by inventor age. The gap between younger and older inventors in the propensity to work at small firms is much smaller than for young firms.[14] Table 7 shows the distribution of inventor jobs by firm size groups. About 10 percent of inventor jobs are at firms with 20 or fewer employees while almost 64 percent of inventor jobs are in firms with over 1000 employees. Inventors in the middle of the inventor age distribution, from 26 to 55 years old, are the most likely to work for the largest firms.

Inventor productivity varies significantly by firm age and firm size. Table 8 shows inventor productivity, measured by average citations received per grant, by firm age and firm size groups. Inventors in mid-size (21 to 250) young firms generate the most impactful patents. Inventors at old, small firms tend to have the lowest citations per patent grant. Inventors at very old, very large firms have citations per grant about 40 percent lower than those at mid-sized young firms (6.0 compared to 10.0). Except among the oldest group (21+), there is an inverse-U shape within firm age bins by firm size. That is, citations rise among larger firms, for a given firm age up to a certain point and then decline for the largest firms. Among the oldest firms, however, there is a strictly monotonic rise across firm size bins.

Focusing on super star inventors, those in the top 10 percent of the citations received distribution within an inventor cohort, we find that the share working for young firms declined more among inventors with the most impactful patents.[15] Figure 8 shows

[13]For this exercise we utilize PatentsView's disambiguated inventor and assignee identifiers. As shown in Table 3, the PatentsView inventor identifiers line up well with our fully triangulated PIK matches, an indication that the inventor disambiguation in PatentsView is high quality.

[14]The measure of firm employment is derived from the LBD, which captures full- and part-time employees, including salaried officers and executives of corporations, who are on the payroll in the pay period including March 12. Included are employees on paid sick leave, holidays, and vacations. We use LBD-LEHD linkages developed by Haltiwanger et al. (2014).

[15]Our identification of super star inventors is time varying. To construct this indicator we generate a panel for each inventor that beings with each inventor's first grant. We compute the sum of cumulative 5-year windowed citations received up to and including the given quarter and divide by the mean of the

the change in the share of super star inventors between the first and second half of our sample by the quality of their patent grants. Super star inventors are placed into four equally sized bins based upon the citations received of their patent grants. We see that the share working at young firms declined most for inventors with the most cited patents. There has been a marked flattening of the relationship between quality of inventive output and the likelihood of working at a young firm over this period.

3.3 Earnings

There are a number of reasons inventor earnings might be related to aggregate innovative activity. An individual's innovative effort is responsive to incentives in both corporate and university settings (Lerner and Wulf, 2007; Lach and Schankerman, 2008). Moreover, evidence suggests that inventors employed by firms are able to capture rents associated with patent grants (Toivanen and Väänänen, 2012) and that taxation affects innovative effort (Akcigit et al., 2016, 2021). Firms may choose to share rents associated with a successful innovation in order to limit attrition among its high skill workers (Kline et al., 2019). Higher inventor compensation may also be a way for incumbent firms to attract talent away from entrants (Akcigit and Goldschlag, 2023). The distribution of returns to high skilled labor, such as inventors, may also have implications for overall earnings inequality. Our data allow us to describe the earnings of U.S. inventors, its skewness, and how earnings relates to inventive activity.

We show the earnings distribution among inventors several ways. First, we examine the skewness of earnings among inventors relative to other workers. Figure 9 shows the Lorenz curve for inventors and a sample of all workers, which captures the cumulative earnings of a population of workers. The y-axis of the figure captures the cumulative share of earnings and the x-axis shows the cumulative share of individuals, ordered by earnings. Perfect earnings equality would place the curves on the 45 degree line. Similar to previous analyses, super star inventors are those in the top 10 percent of the cumulative within-cohort citation distribution. For the bottom 50 percent of individuals and inventors, the Lorenz curve for all workers is below the inventor curves, suggesting the bottom half of all workers accounts for less total earnings for all workers than the bottom half of inventors accounts for total earnings of inventors. The curve for inventors bows out at the top end of the plot, suggesting more high-end earnings inequality among

same measure among all inventors of the same age in that quarter. Using a 5-year window of citations limits the impact of censoring near the end but also constrains the number of years. An inventor is identified as a super star if they are in the top 10% of this quarter-age demeaned cumulative citation measure.

inventors. This is especially true for super star inventors–the bottom 90 percent of super star inventors account for 61 percent of super star inventor earnings. In contrast, the bottom 90 percent of all inventors account for nearly 69 percent of inventor earnings and among all workers that figure is over 71 percent.

Next we provide a more direct comparison of inventor earnings to the earnings of all workers. Table 9 shows the percent of inventors and super star inventors earning more than each percentile of earnings distribution of all workers. Inventors are over represented at the top of the earnings distribution. Over 92 percent of all inventors and almost 98 percent of super star inventors earn more than the 60th percentile of the earnings distribution. The majority of inventors (63%) and the overwhelming majority of super star inventors (88%) are within the top decile of the earnings distribution. At the upper tail of the distribution, the top 1 percent, almost 8 percent of inventors and 19 percent of super start inventors earn more than 99 percent of the population. This suggests that a significant share of individuals at the top of the earnings distribution are inventors.

Consistent with the literature showing that inventor effort is responsive to earnings, we find a strong relationship between earnings and the quality of inventive output. To examine this, we estimate a simple regression of earnings on citations received. Earnings is measured as the average within a window around the current quarter to limit the impact of transient earnings shocks around the time of a patent grant. We segment inventor jobs into deciles of the windowed earnings distribution and control for age, along with year, industry, firm age, and firm size fixed effects. Estimates are shown in Figure 10. We find a positive and monotonic relationship between earnings and citation weighted patent counts. Relative to inventors in the bottom earnings decile, inventors in the top decile produce patents with over 50 percent more citations. Though causality certainly runs in both directions, these partial correlations are consistent with the literature showing the financial returns to inventors of their innovative activity (Toivanen and Väänänen, 2012; Kline et al., 2019).

3.4 Employment Dynamics

Numerous measures of employment dynamics exhibit trend declines in recent decades. Declines appear in measures of job creation and destruction from firm and establishment data, job flows from survey data, or hires, separations, and job-to-job flows from administrative employee employer data (Davis et al., 2006, 2012; Lazear and Spletzer, 2012; Hyatt and Spletzer, 2013). The reallocation of inventors from unproductive firms to productive ones may be especially important for aggregate economic performance and productivity growth. Our data allow us to measure the rate at which inventors change jobs, start new firms, and relocate across geographies.

Inventors saw a significant decline in the hire and separation rates between 2000 and 2016. Figure 11 shows the hire and separation rates for inventors and a sample of all workers weighted by the industry composition of the inventor sample. Hire and separation rates for inventors remained below that of the overall workforce over the entire period. The overall work force also saw a decline in both the rate of hires and separations but both rates stabilize after 2010. Inventors, on the other hand, exhibit a decline in hire rates through the end of the sample period and a separation rate that stabilizes after 2010 but at a much lower level than in the early 2000s.

Another measure of economic dynamism among inventors is the rate at which they start new firms. The rate at which new businesses have started in the U.S. has declined for decades (Decker et al., 2014). One might hope that despite this more persistent, general decline in entrepreneurship the number of high quality entrants remains robust (Guzman and Stern, 2020). Indeed, Miranda and Zolas (2017) find that among non-employers, inventors are more than twice as likely than non-inventors to transition to employer businesses. We find, however, that inventors are less likely to become entrepreneurs over time. Figure 12 shows estimates of the change in the likelihood of an inventor becoming an entrepreneur.[16] Specifically, the figure shows differences in the rate of entrepreneurship, relative to 2000, for super star inventors and non-super star inventors, controlling for individual fixed effects. The share of inventors being a firm founder each year is quite low. In 2000, about 0.64 percent of inventors started a business, about 1 percent of super star inventors and 0.59 percent of non-super start inventors. The estimates suggest that the already low rate of entrepreneurship among inventors fell significantly. Our estimates imply that by 2015, the rate at which non-super star inventors started businesses fell by 41 percent and for super star inventors the decline was even steeper, falling by 57 percent.

Finally, we explore the geographic distribution of inventors. The geographic distribution of inventive activity may have significant economic consequences if local exposure

[16] Entrepreneurship is measured as in Choi et al. (2021). Inventors are flagged as entrepreneurs based upon when they first have earnings at a startup and how their earnings compare to other workers at the startup. For corporations, an inventor will be flagged as a founder of the startup if they have positive earnings in the first quarter of a firm's operations and are among the top three earning workers in the firm's first year. For sole proprietorships, an inventor will be flagged as an entrepreneur if they appear as the business owner in tax filings or are among the top two earners in the firm's first year. Partnerships are excluded from the firm founder database because business owners are prohibited from paying themselves wages that would appear in the LEHD data. See Choi et al. (2021) for additional measurement details.

to inventors affects an individual's propensity to become an inventor Bell et al. (2019). Consistent with the literature on agglomeration and the geographic clustering if innovative activity, we find geographic concentration of U.S. inventors. Indeed, the top 20 counties by number of inventors accounted for about 39 percent of inventors in 2001 but over 47 percent of inventors in 2016. Figure 13 shows the change in the share of inventors in the 20 largest counties by inventor count over time. The figure shows a similar measure for all workers—the share of all workers accounted for by the 20 largest counties by employment. For inventors, by 2016 the share of all inventors in those top 20 counties rose by nearly 15 percent. We do not see a similar rise in concentration among all workers, which rose modestly from 21.6 percent in 2000 to 22 percent in 2016. These patterns are consistent with papers documenting the rising concentration of patenting activity in large urban and tech centers (Bettencourt et al., 2007; Chattergoon and Kerr, 2022).

Related, we find that inventors are also less likely to change employment across-state boundaries. This might be expected given the decades long decline in mobility rates, a pattern that has been related to declining labor market fluidity (Molloy et al., 2014). Figure 14 shows the share of inventors and the share of LEHD workers that change employment across state boundaries each year. Among all workers, this share rose in the early and mid 2000s, declined during the Great Recession, then rose again from 2009 to 2016. Inventors largely follow a similar pattern through the Great Recession, but did not experience a similar rise in the mid 2010s. Inventors are increasingly less likely to change employment across state lines when compared to the typical worker.

Taken together, these results suggest that the labor market outcomes of inventors are becoming less dynamic over time. Inventors are less likely to change jobs, less likely to start a new firm, increasingly concentrated in relatively few locations, and less likely to change employment across state boundaries.

4 Conclusions

There has long been interest in the characteristics of inventors and their role in the innovation process. Only recently have researchers begun to link inventors in patent documents to large-scale survey, census, and administrative data sources. Those linkages dramatically expand the types of analyses that can be done to open the "black box" of innovative activity. In this paper we describe an effort to link patent inventors to Census microdata on individuals and their employment outcomes. We describe a number of patterns in the data that are relevant to both policy makers and researchers interested in the role of inventors in the economy.

First, examining the demographics of inventors, we find that over 30% of inventors were born outside the U.S., the population of inventors as a whole is aging, African Americans are significantly under represented among inventors and the share of Asians among inventors nearly doubled from 2000 to 2016. We also find, consistent with prior work, that females are also underrepresented among inventors, accounting for less than 12% of inventors. Combining information on demographics and employment outcomes allows us to examine how these patterns of representation vary by sector, with female inventors being more common in the Health Care, Social Assistance, and Education sectors while foreign born inventors tend to be employed in the Information, Education, and Professional and Scientific Services sectors.

Linkages to employee-employer data allow us to characterize the types of firms that employ inventors. We find that inventors are concentrated in older, larger firms. Almost 63% of inventors work in firms with more than 1,000 employees that are over 20 years old. Inventors at young firms tend to produce more impactful patents, but highly productive inventors are increasingly less likely to work in young firms over time. We also find that inventor earnings is highly skewed, with 63% of inventors falling in the top 10% of the earnings distribution and 19 percent of super star inventors are within the top 1 percent of the earnings distribution. Earnings also appears to track closely inventor output in terms of citation weighted patent grants.

Finally, we find declining employment dynamism among inventors. The hire and separation rate among inventors fell from about 6% and 7% in 2000 to 4% in 2016. Inventors are less likely to become entrepreneurs over time, and they are becoming increasingly geographically concentrated. The share of inventors working in the 20 counties with the most inventors rose from 39% to 47% over this period. Moreover, inventors appear less likely to move across state lines over time.

Observing these patterns in inventor characteristics and employment dynamics is made possible by matching the inventors listed on patent documents to the survey, census, and administrative data housed at the Census Bureau. These linkages between inventor records and person identifiers within Census microdata will be made available to researchers with approved access via the Federal Statistical Research Data Center (FSRDC) system.[17] We hope that by making these links available to qualified researchers on approved projects that the scope of questions that can be asked and answered about inventors and innovative activity will increase considerably.

[17]See `https://www.census.gov/about/adrm/fsrdc.html` for details about the FSRDC system and how to apply for access to confidential data (accessed 09/2022).

References

Abowd, John M., L. McKinney, Kevin, and Nellie L. Zhao (2017) "Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data,"Technical report, Center for Economic Studies CES-17-24.

Aghion, Philippe, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen (2017) "The Social Origins of Inventors," National Bureau of Economic Research Working Paper 24110.

——— (2018) "On the returns to invention within firms: Evidence from Finland," in *AEA Papers and Proceedings*, 108, 208–12.

——— (2023) "Parental Education and Invention: The Finnish Enigma," *International Economic Review*, forthcoming.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva (2016) "Taxation and the international mobility of inventors," *American Economic Review*, 106 (10), 2930–81.

Akcigit, Ufuk and Nathan Goldschlag (2023) "Where Have All the "Creative Talents" Gone? Employment Dynamics of US Inventors,"Technical report.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas (2017a) "Immigration and the Rise of American Ingenuity," *American Economic Review, Papers and Proceedings*, 107 (5), 327–331.

——— (2017b) "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," National Bureau of Economic Research Working Paper 23137.

Akcigit, Ufuk, John Grigsby, Tom Nicholas, and Stefanie Stantcheva (2021) "Taxation and Innovation in the Twentieth Century," *Quarterly Journal of Economics*, 137 (1), 329–385.

Akcigit, Ufuk, Jeremy G Pearce, and Marta Prato (2020) "Tapping into Talent: Coupling Education and Innovation Policies for Economic Growth," National Bureau of Economic Research Working Paper 27862.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen (2019) "Who Becomes an Inventor in America? The Importance of Exposure to Innovation," *Quarterly Journal of Economics*, 134 (2), 647–713.

Bettencourt, Luis MA, Jose Lobo, and Deborah Strumsky (2007) "Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size," *Research policy*, 36 (1), 107–120.

Breschi, Stefano, Francesco Lissoni, and Gianluca Tarasconi (2017) "Inventor data for research on migration and innovation: The ethnic-inv pilot database," in *The International Mobility of Talent and Innovation: New Evidence and Policy Implications*: Cambridge University Press.

Cao, Chunfang, Xiaohui Li, Xiaoyang Li, Cheng Zeng, and Xuan Zhou (2021) "Diversity and inclusion: Evidence from corporate inventors," *Journal of Empirical Finance*, 64, 295–316.

Celik, Murat Alp (2023) "Does the Cream Always Rise to the Top? The Misallocation of Talent and Innovation," *Journal of Monetary Economics*, 133, 105–128.

Chattergoon, Brad and William R Kerr (2022) "Winner takes all? Tech clusters, population centers, and the spatial transformation of US invention," *Research Policy*, 51 (2), 104418.

Choi, Joonkyu, Nathan Goldschlag, John C Haltiwanger, and J Daniel Kim (2021) "Founding Teams and Startup Performance," National Bureau of Economic Research Working Paper 28417.

Cook, Lisa D and Chaleampong Kongcharoen (2010) "The Idea Gap in Pink and Black," National Bureau of Economic Research Working Paper Nr. 16331.

Davis, Steven J, R Jason Faberman, and John Haltiwanger (2012) "Labor market flows in the cross section and over time," *Journal of Monetary Economics*, 59 (1), 1–18.

Davis, Steven J, John Haltiwanger, Ron Jarmin, and Javier Miranda (2006) "Volatility and dispersion in business growth rates: Publicly traded versus privately held firms," in Acemoglu, Daron, Kenneth Rogoff, and Michael Woodford eds. *NBER macroeconomics annual*, 107–179: Mit Press.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda (2014) "The Role of Entrepreneurship in U.S. Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, 28 (3), 3–24.

Dreisigmeyer, David, Nathan Goldschlag, Marina Krylova, Wei Ouyang, and Elisibeth Perlman (2018) "Building a Better Bridge: Improving Patent Assignee-Firm Links,"Technical report, Center for Economic Studies CES-TN-2018-01.

Foster, Lucia, Cheryl Grim, and Nikolas Zolas (2020) "A portrait of US firms that invest in R&D," *Economics of Innovation and New Technology*, 29 (1), 89–111.

Graham, Stuart JH, Cheryl Grim, Tariqul Islam, Alan C Marco, and Javier Miranda (2018) "Business Dynamics of Innovating Firms: Linking US Patents with Administrative Data on Workers and Firms," *Journal of Economics & Management Strategy*, 27 (3), 372–402.

Griliches, Zvi (1992) "The search for R&D spillovers," *Scandinavian Journal of Economics*, 94, 29–47.

——— (1998) "Patent statistics as economic indicators: a survey," in *R&D and productivity: the econometric evidence*, 287–343: University of Chicago Press.

Guzman, Jorge and Scott Stern (2020) "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014," *American Economic Journal: Economic Policy*, 12 (4), 212–43.

Hall, Bronwyn (1996) "The private and social returns to research and development," in *Technology, R&D, and the Economy*, 140–183: Brookings Institution and the American Enterprise Institute.

Haltiwanger, John, Henry R Hyatt, Erika McEntarfer, Liliana Sousa, and Stephen Tibbets (2014) "Firm Age and Size in the Longitudinal Employer-household Dynamics Data,"Technical report, Center for Economic Studies CES-WP-14-16.

Henderson, Cody and Henry Hyatt (2012) "Estimation of Job-to-Job Flow Rates Under Partially Missing Geography,"Technical report, Center for Economic Studies CES-12-29.

Hyatt, Henry R and James R Spletzer (2013) "The recent decline in employment dynamics," *IZA Journal of Labor Economics*, 2 (1), 5.

Jaravel, Xavier, Neviana Petkova, and Alex Bell (2018) "Team-specific capital and innovation," *American Economic Review*, 108 (4-5), 1034–73.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar (2019) "Who profits from patents? rent-sharing at innovative firms," *Quarterly Journal of Economics*, 134 (3), 1343–1404.

Lach, Saul and Mark Schankerman (2008) "Incentives and invention in universities," *RAND Journal of Economics*, 39 (2), 403–433.

Lamoreaux, Naomi R and Kenneth L Sokoloff (1999) "Inventive activity and the market for technology in the United States, 1840-1920,"Technical report, National Bureau of Economic Research Working Paper 7107.

Lazear, Edward P and James R Spletzer (2012) "Hiring, churn, and the business cycle," *American Economic Review*, 102 (3), 575–79.

Lerner, Josh and Julie Wulf (2007) "Innovation and incentives: Evidence from corporate R&D," *the Review of Economics and Statistics*, 89 (4), 634–644.

Marco, Alan C, Joshua D Sarnoff, and Charles deGrazia (2019) "Patent Claims and Patent Scope," *Research Policy*, 48 (9), 103790.

Miranda, Javier and Nikolas Zolas (2017) "Measuring the impact of household innovation using administrative data," in *Measuring and Accounting for Innovation in the 21st Century*: University of Chicago Press.

Molloy, Raven, Christopher L Smith, and Abigail K Wozniak (2014) "Declining migration within the US: The role of the labor market,"Technical report, National Bureau of Economic Research, WP-20065.

Monath, Nicholas, Christina Jones, and Sarvo Madhavan (2021) "Patents View: Disambiguating Patent Inventors, Assignees, and their Locations,"Technical report, American Institute for Research.

Ouimet, Paige and Rebecca Zarutskie (2014) "Who works for startups? The relation between firm age, employee age, and growth," *Journal of financial Economics*, 112 (3), 386–407.

Schmookler, Jacob (1957) "Inventors past and present," *The Review of Economics and Statistics*, 39 (3), 321–333.

Toivanen, Otto and Lotta Väänänen (2012) "Returns to inventors," *Review of Economics and Statistics*, 94 (4), 1173–1190.

USPTO (2019) "Progress and Potential: A profile of women inventors on U.S. patents,"Technical report, USPTO.

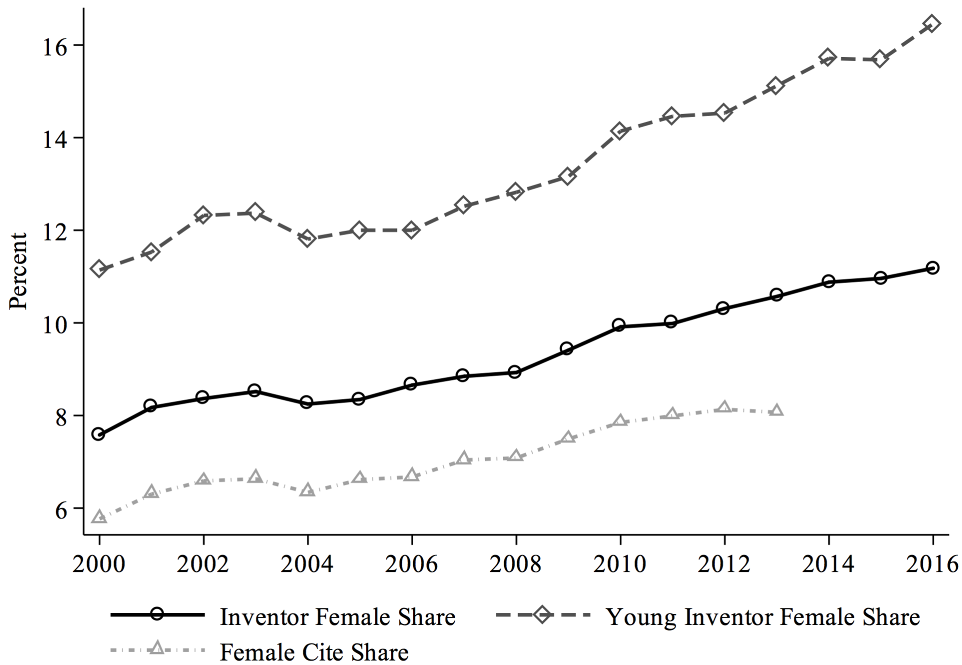
Vilhuber, Lars (2018) "LEHD Infrastructure S2014 files in the FSRDC,"Technical report, Center for Economic Studies CES-WP-2018-27.

Wagner, Deborah and Mary Lane (2014) "The person identification validation system (PVS): applying the Center for Administrative Records Research and Applications' (CARRA) record linkage software,"Technical report, Center for Economic Studies, US Census Bureau, WP-2014-01.

Winston, Sanford (1937) "Bio-social characteristics of American inventors," *American Sociological Review*, 2 (6), 837–849.

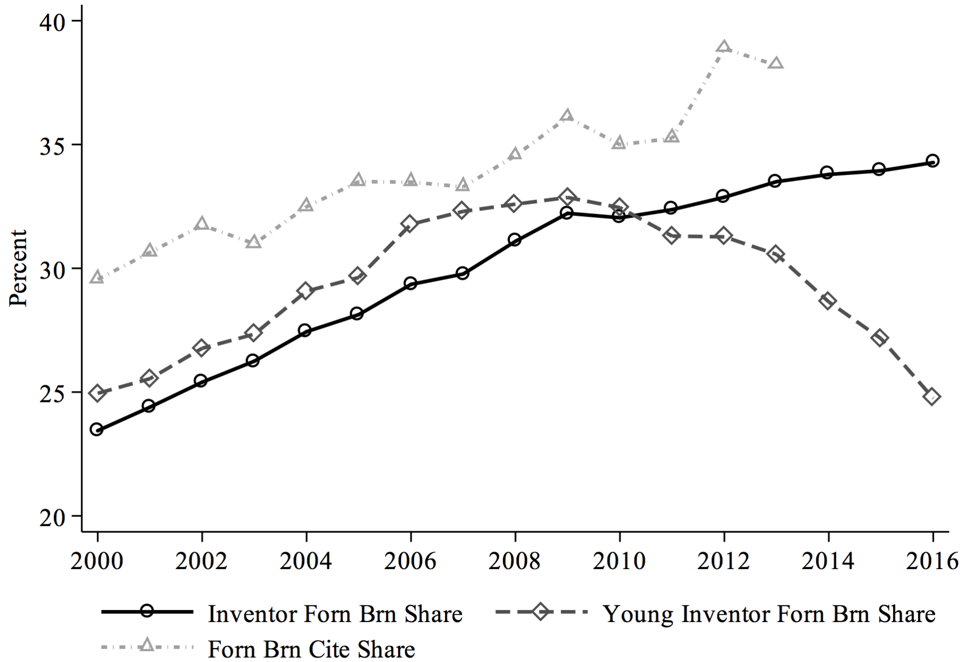
Figures

FIGURE 1: INVENTOR FEMALE SHARE



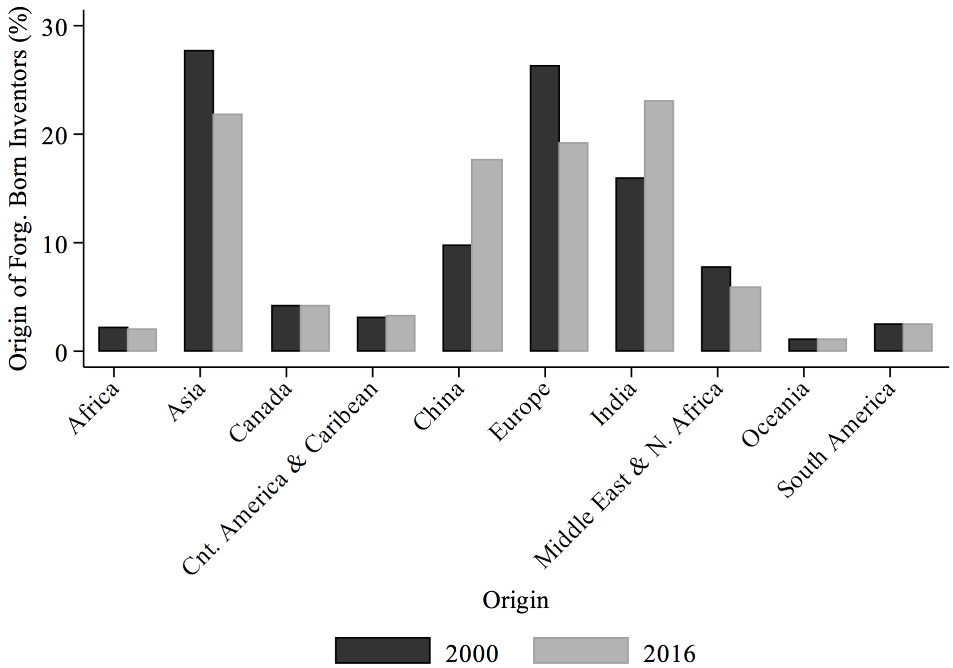
Source: Inventor Employment History
Notes: Female share is the share of grant positive inventors that are female. Young inventor female share is the share of inventors ≤ 35 that are female. Female cite share is the share of cites associated with female inventors. Shares are averaged across quarters within each year.

Figure 2: Inventor Foreign Born Share



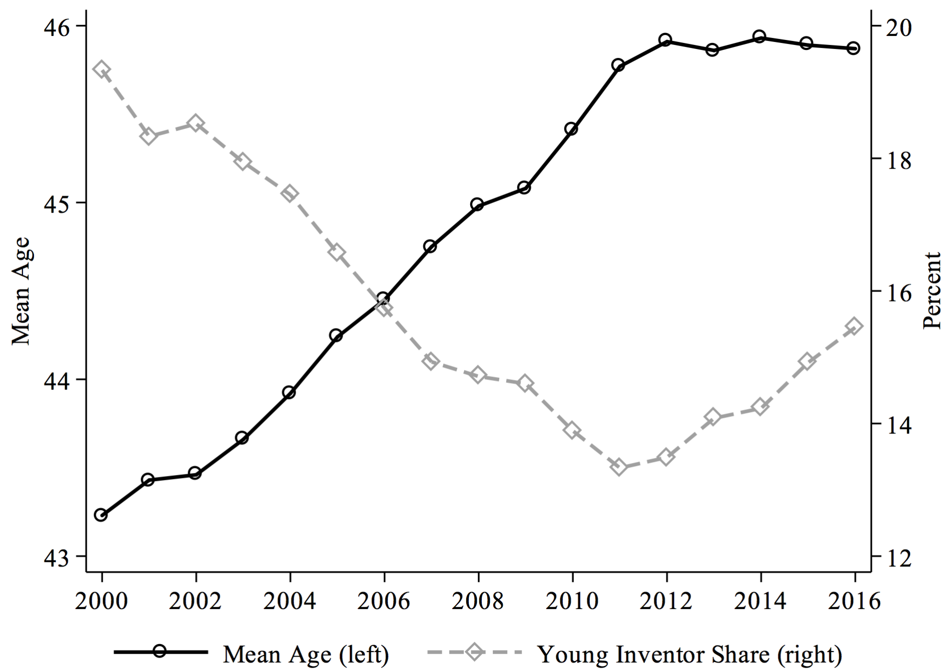
Source: Inventor Employment History
Notes: Foreign born share is the share of grant positive inventors that are foreign born. Young inventor foreign born share is the share of inventors ≤ 35 that are foreign born. foreign born cite share is the share of cites associated with foreign born inventors.

Figure 3: Origin of Foreign Born Inventors



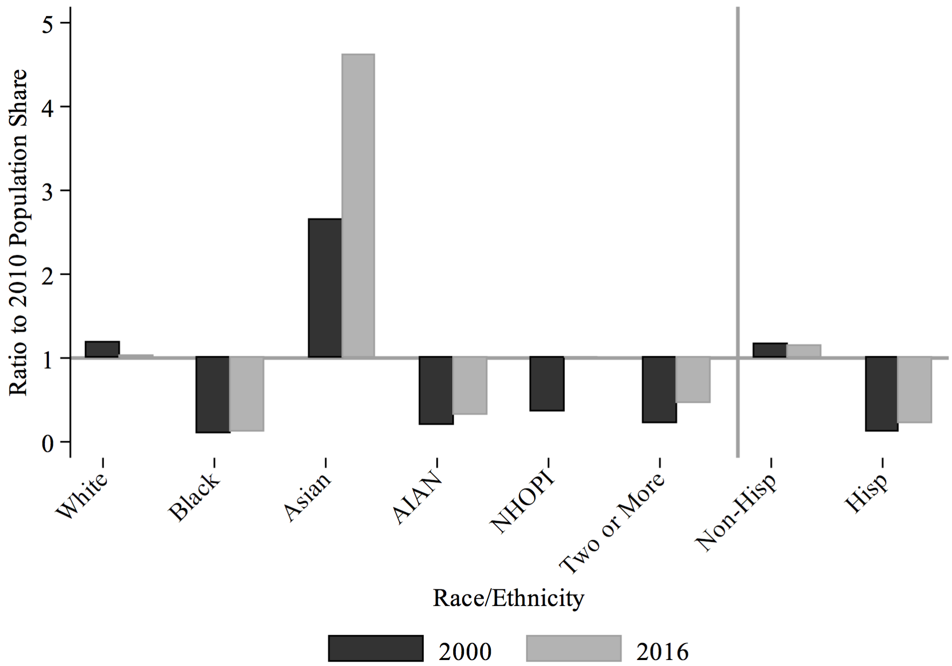
Source: Inventor Employment History
Notes: Share of grant positive foreign born inventors by origin region/country for 2000 and 2016. Country and region groups are recoded combinations of the Place of Birth codes outlined in Vilhuber (2018). Asia, for example, is all Asian countries excluding China and India. Specifically, Asian includes [1,2,C,D,M,N,P]. Europe includes [F,J,U,X,Y,Z]. Central America & Carribean includes [4,5,B,G,H,I,O,Q,S,W]. South America includes [6,T]. Middle East & Africa includes [3,V].

Figure 4: Inventor Age



Source: Inventor Employment History
Notes: Inventor mean age is the mean age of all grant active inventors. Young inventor share is the share of grant active inventors ≤ 35.

Figure 5: Inventor Race and Ethnicity Composition Relative to 2010 Population

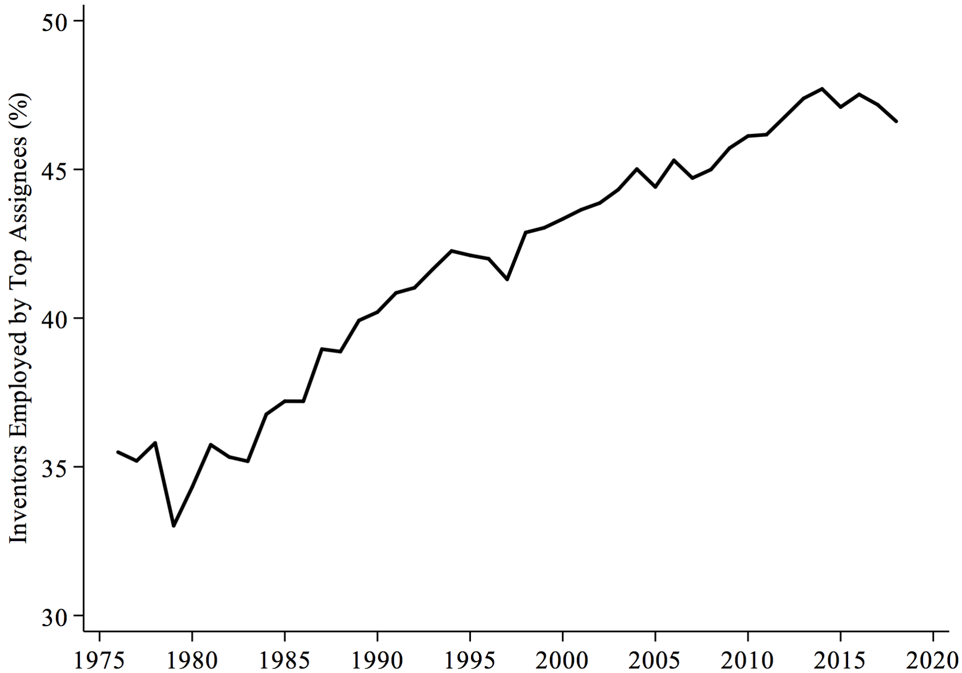


Source: Inventor Employment History, 2010 Decennial Census
Notes: Ratio of inventor population share, among grant active inventors, to 2010 population share per the 2010 Decennial Census. Figure shows the ratio using the 2000 and 2016 inventor race and ethnicity shares. AIAN stands for American Indian or Alaska Native and NHOPI stands for Native Hawaiian or Other Pacific Islander.

Figure 6: Share of Inventors at Young Firms

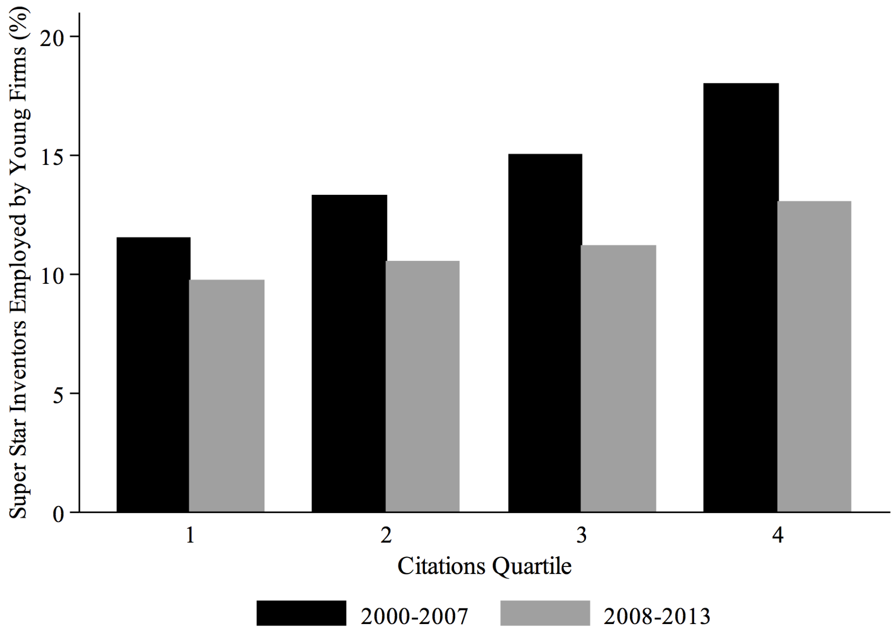
Source: Inventor Employment History
Notes: Figure shows the percent of inventors at young firms ($\leq$ 5 years old).

Figure 7: Share of Inventors on Grants with Assignees with Many Inventors



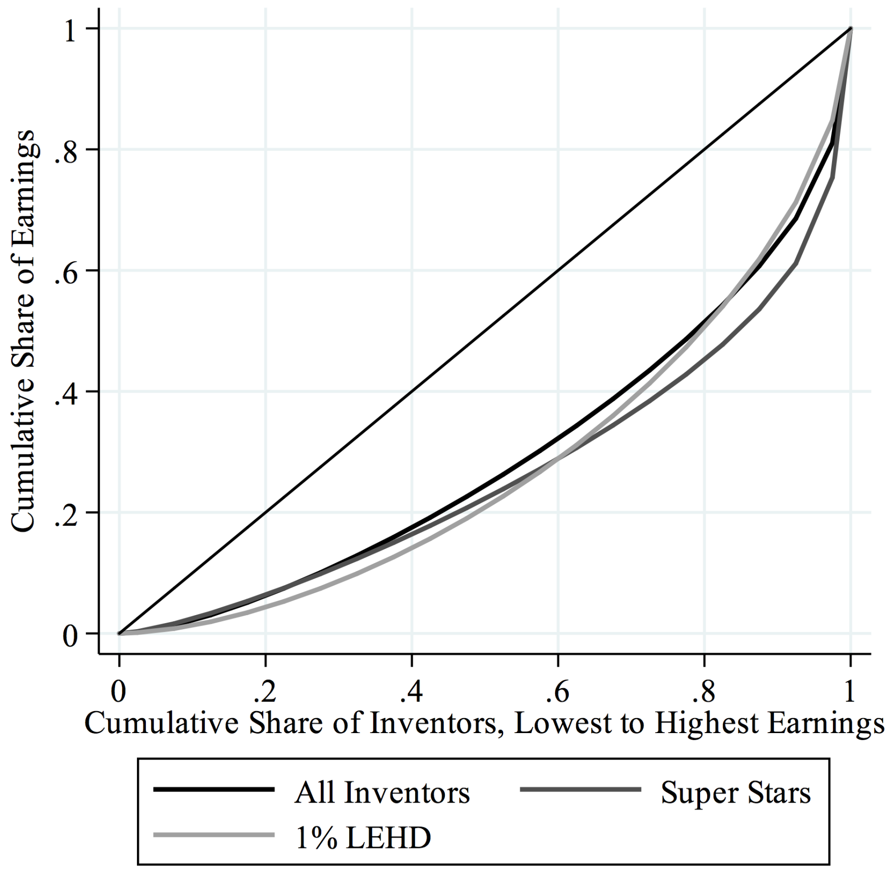
Source: PatentsView
Notes: Figure shows the percent of inventors (disambiguated using PatentsView inventor IDs) on patent grants each year that were linked, via a patent grant, to an assignee (disambiguated using PatentsView assignee IDs) that itself is among the top one percent of assignees by the number of associated unique inventors. The number of associated inventors necessary to be classified as among the top one percent is 81 in 1980 and rises to 154 in 2018.

Figure 8: Super Star Inventors at Young Firms, Pre & Post 2008



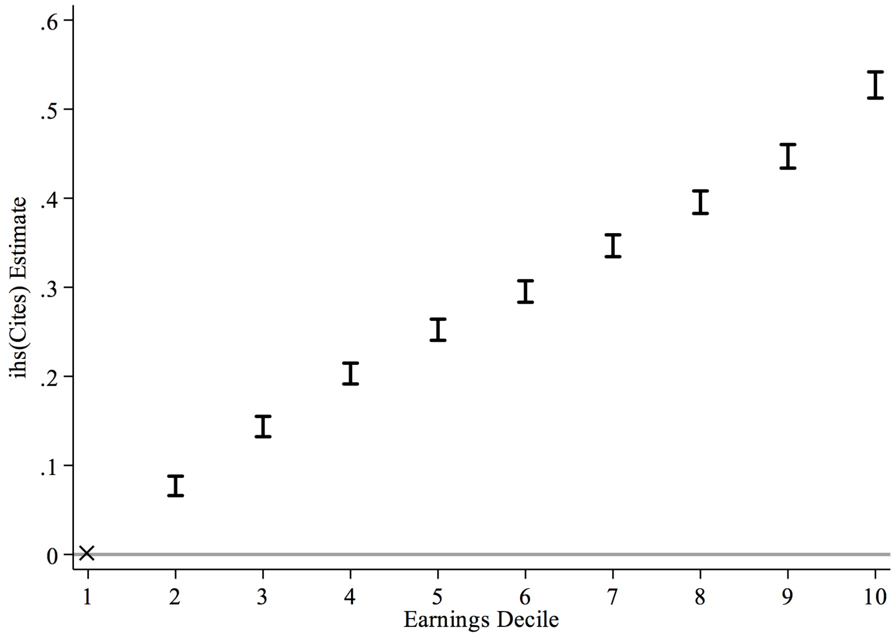
Source: Inventor Employment History
Notes: Figure shows the percent of Super Star inventor jobs at young firms ($\leq$ 5 y.o.) in the first and second half of the sample by citation quartiles. Super star inventors are those in the top 10 percent of the cumulative citations received distribution within cohort. Patent citations received quartiles split the sample of super star inventors into four groups based upon the citations received of their granted patents.

Figure 9: Inventor Earnings Lorenz



Source: Inventor Employment History
Notes: Figure shows the proportion of earnings assumed by the bottom x percent of the people. For the bottom x percent of inventors, what percentage (y percent) of the total earnings they have. The percentage of inventors is plotted on the x-axis, the percentage of earnings on the y-axis.

Figure 10: Earnings and Citations

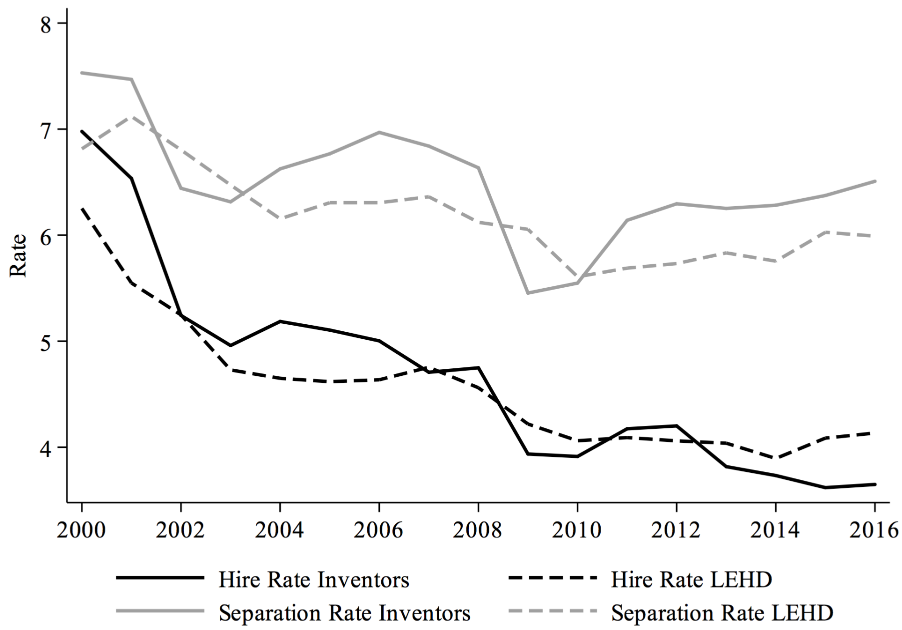


Source: Inventor Employment History
Notes: This figure shows estimates of λ_j, with $j = 1$ being the excluded group, from the following equation:

$$ihs(Cites_{i,t}) = \alpha + \sum_{j=1}^{10} \lambda_j EarnL2F2[j]_{i,t} + \boldsymbol{X} + \epsilon_{i,t} \quad (1)$$

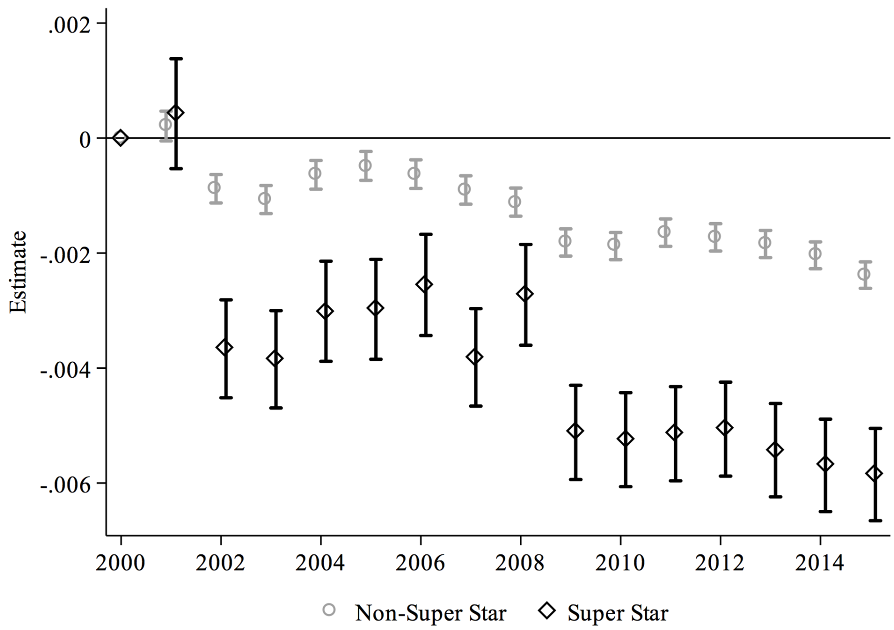
ihsCites is the ihs() transformed 5-year window cites of patents granted in t. $EarnL2F2$ is the decile bin of average earnings in $t-2$ to $t+2$. $\boldsymbol{X}$ contains age and age squared, year FE, 4-digit NAICS FE, firm age group FE, and firm size group FE. Clustered at the individual level.

Figure 11: Inventor Hire and Separation Rates



Source: Inventor Employment History
Notes: Figure shows Hire and Separation (QWI definition of stable hires and separations) for inventors and a 1% sample of the LEHD for comparison. The 1% LEHD sample is weighted to reflect the time varying industry composition of the inventor sample and thus captures the hire and separation of workers in industries in which inventors are most frequent.

Figure 12: Change in Inventor Entrepreneurship Rate

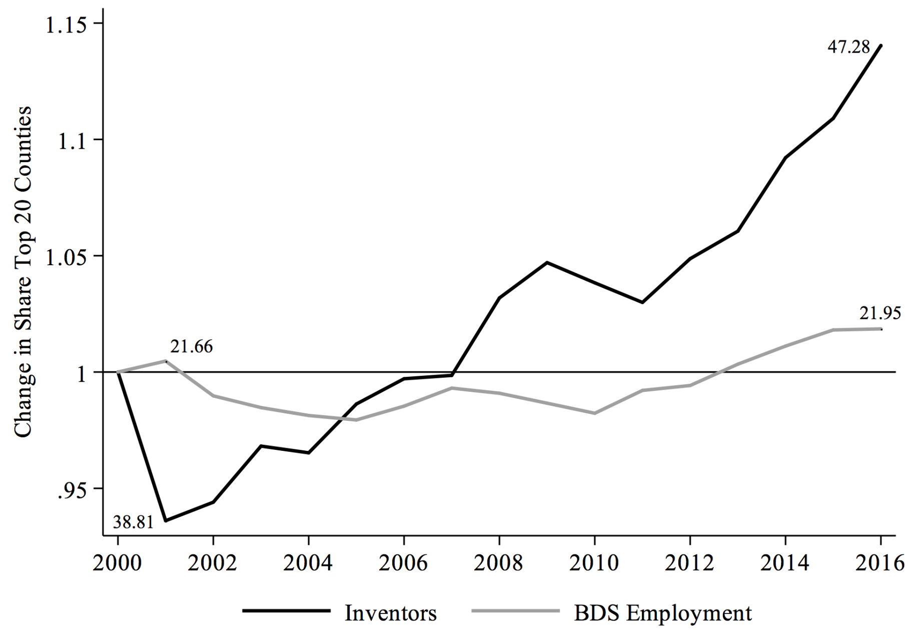


Source: Inventor Employment History
Notes: This figure shows estimates from the following:

$$Entrep_{i,t} = \alpha + \sum_{k=2001}^{2015} \beta_k D^k + \psi_i + \epsilon_{i,t} \tag{2}$$

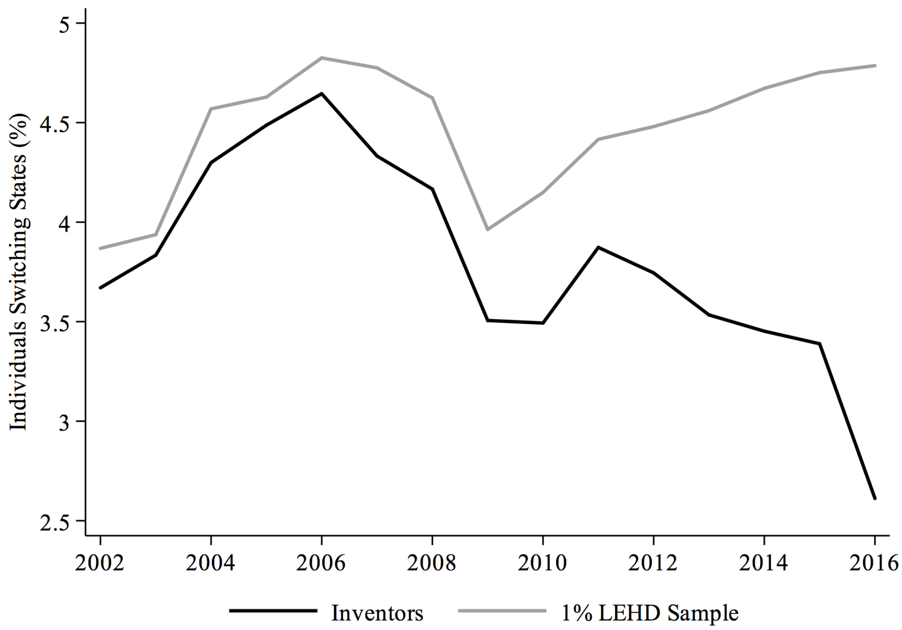
Where $Entrep_{i,t}$ is an indicator of whether an inventor i is a startup founder in year t. D^k are year dummies, and ψ_i are inventor fixed effects. For corporations, an inventor will be flagged as a founder of the startup if they have positive earnings in the first quarter of a firm's operations and are among the top three earning workers in the firm's first year. For sole proprietorships, an inventor will be flagged as an entrepreneur if they appear as the business owner in tax filings or are among the top two earners in the firm's first year. Partnerships are excluded from the firm founder database because business owners are prohibited from paying themselves wages that would appear in the LEHD data. See Choi et al. (2021) for additional measurement details.

Figure 13: Change in Top 20 County Concentration



Source: Inventor Employment History, Business Dynamics Statistics (2018)
Notes: Figure shows change in share of grant active inventors by county accounted for by the top 20 counties by the count of grant active inventors. Change is measured as the ratio shr_t/shr_{2000}. Note that inventors are allowed to contribute to multiple counties in a year. The BDS employment change is measured analogously as the share of employment accounted for by the 20 largest counties by employment, which are not necessarily the same counties as the top counties for inventors. The set of counties is time varying, though there is relatively little volatility in the set of counties over time. The labeled points show the ending share in 2016.

Figure 14: State Movers



Source: Inventor Employment History
Notes: Share of all (not just grant active) inventors and 1 percent LEHD sample workers that changed states at least once for their dominant BoQ job in a year.

Tables

Table 1: Inventor Record Match Type

Match Type	Percent of Matches
A1	.62
A2	.04
A3	.01
PatentsView	.08
PVS	.25

Source: Inventor Record-PIK Match, LEHD, PatentsView
Notes: A1, A2, and A3 are triangulation match types in which the PIK is assigned with both EIN and FIRMID matching between the job and BR-matched firm, only FIRMID match, and only EIN match respectively. PatentsView is the share of inventor records matched by using the disambiguated inventor identifiers in the PatentsView database to extend PIK matches to unmatched inventor records. PVS is the share of inventor records matched using unique PVS matches.

Table 2: Matched and Unmatched Inventor Record Characteristics

Variable	Unmatched	Matched
Cites Made	12.51	30.88
Cites Received	3.643	6.51
Cites Received, Demeaned	.7995	1.336
Independent Claim Count	2.557	3.103
Team Size	2.367	2.758

Source: Inventor Record-PIK Match, PatentsView
Notes: Shown are mean patent-level characteristics of unmatched and matched inventor records. Cites Made is the count of citations made. Cites Received, 5 Year Window is the count of citations received within the first five years after the patent grant. Cites Received, Demeaned is the count of citations received demeaned within application year and technology class. Independent Claim Count is the count of independent patent claims as developed by Marco et al. (2019). Team Size is the count of inventors on the patent.

Table 3: Agreement between PatentsView Inventor ID and PIKs

PIK Match Count	InvID Share	Modal PIK Share	InvID Match Count	PIK Share	Modal InvID Share
1	94.47	1	1	96.63	1
2	4.962	.7104	2	2.725	.7204
3	.5	.6232	3	.4005	.6214
4	.05844	.58	4	.1202	.5439
5+	.01181	.5555	5+	.1248	.4558

Source: Inventor Record-PIK Match, PatentsView
Notes: The InvID Share column shows the share of PatentsView inventor IDs that map to some number of PIKs given in the PIK Match Count column. PIK Share shows the share of PIKs that map to some number of InvIDs given in the InvID Match Count column. The Modal PIK Share column shows the share of inventor records accounted for by the most frequent PIK. Similarly, the Modal InvID Share column shows the share of inventor records associated with the most frequent InvID.

Table 4: Change in Race/Ethnicity Composition

Race/Ethnicity	2010 Decennial Share	Inventor Share 2000	Inventor Share 2016
White Alone	72.407	85.26	74.63
Black or African American Alone	12.6089	1.239	1.635
American Indian or Alaska Native Alone	.9497	.1882	.3074
Asian Alone	4.7529	12.58	21.89
Native Hawaiian or Other Pacific Islander Alone	.1749	.06219	.1739
Two or More Races	2.918	.6678	1.363
Hispanic	16.3493	2.002	3.692
Non-Hispanic	83.6507	98	96.31

Source: Inventor Employment History, 2010 Decennial Census
Notes: Inventor shares for 2000 and 2016 are the share of grant active inventors in each year respectively.

Table 5: Inventor Demographics by Sector

Sector	Female	Foreign Born	Age ≤ 35	Age ≥ 56	non-White
Mining, Util., Constr.	5.146	24.3	14.14	20.16	15.22
Manufacturing	7.329	28.55	15.02	14.54	19.72
Wholesale, Retail Trade	9.315	30.22	16.73	13.9	20.99
Transp. & Warehousing	7.942	16.49	15.11	17.31	15.29
Information	8.756	43.45	22.63	7.126	29.61
FIRE	14.32	26.91	17.76	12.23	22.04
Prof., Sci., & Tech. Services	9.489	38.54	15.64	13.47	26.06
Educational Services	14.77	42.64	11.25	23.23	25.74
Health Care & Social Ass.	17.8	31.31	10.96	21.88	20.15
Other	12.39	24.69	14.97	17.34	16.97

Source: Inventor Employment History
Notes: Demographic shares of grant active inventors by sector. Female and Foreign Born are the female and foreign born share respectively. Age ≤ 35 and Age ≥ 56 are the share of grant active inventors aged less than or equal to 35 and greater than or equal to 56 respectively. non-White is the share of non-white inventors. FIRE denotes Finance, Insurance, Real Estate, and Rental and Leasing.

Table 6: Inventor Employer Firm Age Distribution by Inventor Age

Firm Age	All	≤ 25	$26-35$	$36-45$	$46-55$	56+
0 to 5	10.3	14.8	13	10.8	8.6	7.6
6 to 10	7.6	9	8.3	7.9	7	6.7
11 to 20	13.7	16.9	14.9	13.8	12.9	13.1
21+	68.3	59.3	63.7	67.5	71.5	72.7

Source: Inventor Employment History
Notes: Columns show the share of inventor-quarters for a given age group, or among all inventors, across firm age groups.

Table 7: Inventor Employer Firm Size Distribution by Inventor Age

Firm Size	all	≤ 25	$26-35$	$36-45$	$46-55$	$56+$
1 to 20	10.1	13.7	9	9.3	10.3	12.9
21 to 250	16.7	20	16.7	16.3	16.2	17.9
251 to 1000	9.6	9.2	9.6	9.7	9.5	9.7
1000+	63.6	57.1	64.7	64.7	64	59.5

Source: Inventor Employment History
Notes: Columns show the share of inventor-quarters for a given age group, or among all inventors, across firm age groups.

Table 8: Firm Age, Firm Size, and Cites per Grant

	Firm Size			
Firm Age	1 to 20	21 to 250	251 to 1000	1000+
1	8.883	10.02	8.233	7.45
2	6.638	8.762	8.786	7.905
3	5.665	6.617	8.256	7.692
4	4.951	5.033	5.713	6.036

Source: Inventor Employment History
Notes: Average citations per grant received within the first five years after grant by firm age and firm size groups averaged over time. Fage groups: 0 to 5, 5 to 10, 10 to 20, 20+ and LC. Firm size groups: [1, 20], (20 to 250], (250,1000], 1000+.

Table 9: Inventor Relative Wage Distribution

Population Earnings Percentile	Percent of Inventors Above	Percent of Super Star Inventors Above
99	7.71	18.77
98	15.69	37.27
96	32.03	64.5
94	45.3	77.87
92	55.26	84.6
90	62.69	88.35
80	81.47	95.225
60	92.347	97.898

Source: Inventor Employment History
Notes: Table shows the percent of inventors with earnings less than each percentile of earnings distribution for the population. The population is approximated using a 1 percent sample of the LEHD data with a constant state composition. Earnings percentiles and the share of inventors with earnings less than each percentile are calculated within each year-quarter from 2000q1 to 2016q3. This table shows the mean shares across quarters.