The Rise of American Ingenuity: Innovation and Inventors of the Golden Age*

Ufuk Akcigit John Grigsby Tom Nicholas†

June 7, 2017

Abstract

We examine the golden age of U.S. innovation by undertaking a major data collection exercise linking inventors from historical U.S. patents to Federal Censuses between 1880 and 1940 and to regional economic aggregates. We provide a theoretical framework to motivate the micro and macro-level stylized facts we uncover in the data. We show that inventors were highly educated and that father's income and education were important intergenerational transmission channels. Inventors tended to migrate to places that were conducive to innovation and they were positively selected through exit early in their careers. New inventors received more patent citations than incumbent inventors, suggesting a cycle of creative destruction. The financial returns to technological development were high. At the macro-level we identify a strong relationship between patented inventions and long-run economic growth, and use an instrumental variables approach exploiting an historical shift in innovation activity during World War II to show that this relationship could be causal. Finally, we document a U-shaped relationship between top income inequality and innovation, yet innovative places tended to be more socially mobile. Our new data help to address important questions related to innovation and long-run growth dynamics.

Keywords: Economic development, growth, selection, competition, firm dynamics, management, entrepreneurship, creative destruction.

JEL Classifications: O31, O38, O40

*We would like to thank the seminar and conference participants at the NBER Summer Institute Economic Growth (EFJK) Group, Development of the American Economy Group, AEA Meetings, Firms in Macroeconomics Conference, Harvard University, University of Chicago, University of Maryland, University of Montreal, Boston University, Arizona State University, College de France, the USPTO, the University of Houston, the Fuji Conference in Tokyo, Queen's University, Koc University, and Northwestern for very helpful feedback and discussions. We thank the Minnesota Population Center for access to the Census data and Price Fishback, Rick Hornbeck, William Kerr, and Emmanuel Saez for sharing various datasets with us. Akcigit gratefully acknowledges the National Science Foundation, the Alfred P. Sloan Foundation, and the Ewing Marion Kauffman Foundation for financial support. Nicholas received funding from the Division of Research and Faculty Development at Harvard Business School.

†Authors' Addresses: Akcigit: University of Chicago, NBER, and CEPR (uakcigit@uchicago.edu). Grigsby: University of Chicago (jgrigsby@uchicago.edu). Nicholas: Harvard Business School (tnicholas@hbs.edu).

1 Introduction

At the heart of the 25-year old endogenous growth literature is the premise that innovation and technological progress are engines of long-run economic growth (e.g., Romer (1990), Aghion and Howitt (1992)). An abundance of modern data has helped to inform theoretical perspectives on key aspects of the growth process including the impact of firm entry and exit, reallocation, the role of cities, the distribution of talent, and the relationship between inequality, social mobility, innovation and growth (e.g., Duranton and Puga (2001), Banerjee and Duflo (2003), Klette and Kortum (2004), Foster, Haltiwanger, and Syverson (2008), Bloom, Schankerman, and Van Reenen (2013), Acemoglu et al. (2013), Hsieh et al. (2013), Hsieh and Klenow (2014), Jones and Kim (2014), Aghion et al. (2015a), Akcigit and Kerr (2017)). Due to data limitations, however, evidence from longer time horizons remains elusive, despite the large influence that innovations from history such as light bulbs, air conditioners and storage batteries exert on modern society. Studying the creators of these inventions has the potential to shed new light on the innovation and growth literatures.

This paper develops a number of facts about the environment in which many of the essential technologies used today were created, the life cycle of inventors who developed these inventions, and how innovation relates to long run growth, inequality and social mobility. We generate a novel data series matching inventors of patents in the United States to Federal Censuses from 1880 to 1940. Typically such data has only been available for broader populations in modern time periods (Aghion et al. (2015b), Bell et al. (2015)) or historically for specific sub-samples of inventors (e.g., Lamoreaux and Sokoloff (1999)). The new data allow us to examine who became an inventor and the types of environments that were most conducive to innovation and long run growth.

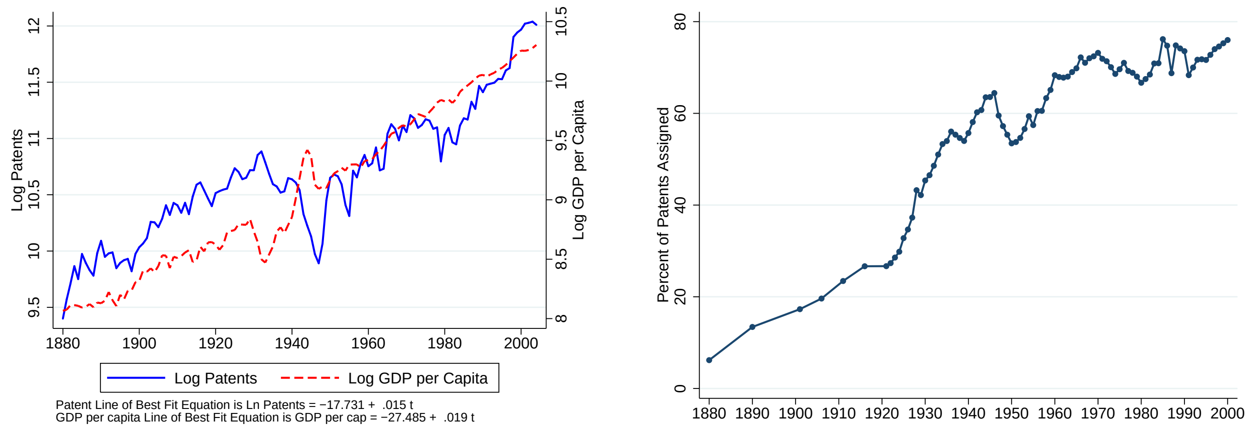
Developing new facts about economic growth has a long tradition going back to Kaldor (1961) who presented six stylized facts around which the theory of economic growth developed. Jones and Romer (2010) updated Kaldor's facts to reflect the subsequent fifty years of data, providing the empirical foundations for modern growth theory. Both papers have facilitated informed discussion and permitted key breakthroughs in our understanding of economic growth. Klette and Kortum (2004) and Banerjee and Duflo (2005) emphasized the discovery of new empirical facts as being important for growth theory to progress. Our paper attempts to establish the fundamental facts regarding the process of innovation during a critical period of U.S. economic development.

The time period we cover is central to recent debates on innovation and growth. We analyze the years that Gordon (2016) associates with the second industrial revolution, which produced major innovations like electricity and the motor vehicle. As an overview of the underlying innovation data, Figure 1 (Panel A) plots the time-series of log patents filed at the USPTO. It shows that innovative activity (proxy measured by patenting) has been growing over time. In keeping with the predictions of the large theoretical literature highlighting the central role of technological progress in endogenous growth we find a positive association between our innovation measures and output growth over the long run.

Panel B shows that most U.S.-based inventors in 1880 developed technologies outside the

Figure 1: Long-Run History of Total Patents Filed in the USPTO

Panel A: Total Patents   Panel B: Share of Corporate Patents



Notes: Panel A: Excluding the exceptional years around World War II (1930-1950), a regression of annualized 5-year GDP per capita growth rates on log patents yields a coefficient of 0.639, which is statistically different from 0 at the 10% level. The correlation between these five year growth rates and log patents is 0.392. Correlation between log patents and log GDP per capita (levels) is 0.890, while the estimated linear relationship between the two is 0.751 (with heteroskedasticity robust standard error 0.0325). Panel B: The share of U.S. patents assigned at their grant date proxies for the amount of corporate patenting. Sources: USPTO, Maddison, Bureau of Economic Analysis, Klein (2013).

boundaries of firms. Over the subsequent 120 years, the share of patents assigned to corporations rose substantially, reflecting the development of R&D labs inside the modern corporation. Given the high share of unassigned patents in the historical data, the innovation process may be well-understood through the life cycle of inventors.

Several famous case studies can motivate our approach. Born to a poor family in rural Ohio, Thomas Edison (1847-1931) faced tight financing constraints in his early career. He ultimately relocated to New Jersey, building the Menlo Park Lab in 1876, a pioneering research laboratory where creative inventors could collaborate to develop new technical ideas. To develop his technologies further Edison accessed capital from a group of financiers, including J.P. Morgan. The investment bank Drexel, Morgan & Co. (which later became J.P. Morgan & Co.) provided loans, acted as a financial intermediary for Edison's firm, and provided wealth management. Edison was granted 1,093 U.S. patents, accruing great wealth in the process. His experience suggests the importance of access to capital, population density, and human interactions in the innovation process. Edison's career also exemplifies the potential for strong financial returns to innovation, and its possible link to social mobility.

Nikola Tesla (1856-1943), a Serbian immigrant, demonstrates the contribution of international migrants to U.S. technological progress. After arriving in America in 1884 at age 28, he began work at the Edison Machine Works in New York. Tesla's career highlights how productivity could change over an inventor's life cycle: he was granted 64 of his 112 U.S. patents (57% of the total) within a decade of his arrival. However, his patenting rate attenuated sharply, and he acquired just 16 patents (14% of the total) after age 45. Tesla believed that private relationships detracted from productive research time. He never built a family. His decision not to marry shows that inventors faced trade-offs when it comes to time allocation.

Finally, Melvin De Groote (1895-1963), one of the most prolific inventors in U.S. history, received two degrees in Chemical Engineering. His highly-educated background was crucial for turning his creativity into valuable innovations. De Groote was granted 925 U.S. patents, mostly developing novel methods to separate crude oil emulsion into its oil and water components. De Groote moved from his state of birth in West Virginia to various places in the U.S. where innovative firms were located.[1]

The above case studies hint that myriad factors, such as immigration, social mobility, access to capital, human interactions, education, and time allocation, might spur innovation. Our new dataset can systematically document these patterns. To organize our exposition, we develop a simple model of innovation using key insights from endogenous growth theory. The model guides our empirics to the relevant correlations in the data, and aids interpretation of the results.

As a broad summary of our approach, we examine the basic demographic facts of inventors: their education, migration decisions, life cycle and the private rewards to innovation by studying the wages of successful inventors from labor income data in the 1940 Census (subject to the limitations associated with the income data from the 1940 Census which we discuss in Section 2). Next, we establish a link between innovation and economic growth at the state level. Finally, we investigate the societal consequences of innovation by establishing the correlation between patenting activity and income inequality or social mobility at the state level.

Empirical Facts

Our analysis uncovers the following stylized facts about innovation:

Fact 1. Inventors were **more educated** on average and were most productive between the age of **36 and 55**.

Fact 2. **Father's income** and **father's education** were highly correlated with becoming an inventor, especially through the effect on the level of a **child's education**.

Fact 3. Inventors were more likely to **have migrated** from their state of birth. They moved to states that were more conducive to innovation.

Fact 4. Inventors were **positively selected** through exit early in their careers, increasing the average productivity (conditional of survival) of a cohort of inventors. However, they were less productive and more likely to exit late in their careers.

Fact 5. The patents of **new inventors received more citations** on average, and were more likely to be in the top decile of the citations distribution.

Fact 6. Successful patentees had substantially higher **labor income**, and produced higher **quality inventions**. For younger inventors, future productivity predicts current income whereas for older inventors, income is predicted by both past and future productivity.

Fact 7. More inventive states **grew faster** on average.

Fact 8. Broad measures of **income inequality**, such as the 90/10 ratio and Gini coefficient, were negatively correlated with innovation at the state level, however, the **top-1% income share** had a U-shaped relationship with innovation.

Fact 9. Innovation was strongly positively correlated with **social mobility**.

[1] De Groote was a consultant or employee to numerous corporations including Hachmeister Lind Chemical of Pittsburgh, Procter and Gamble in Cincinnati, and Petrolite Corporation in Webster Groves, Missouri where he was Vice President and Director of Research.

More specifically, our analysis shows that inventors were not uneducated amateurs; rather they were typically highly-educated individuals who were most productive between the age of 35 and 55. In turn, access to education can be related to the family background of inventors because those with high-earning, or more educated, father's were more likely to become inventors themselves. Equally, while we find that some inventors gained privileged access into a career as an inventor, this effect operated only in the upper tail of the father income distribution. On the intensive margin we find no effect of father's income on patent productivity or quality.

To investigate personal characteristics further, we examine migration patterns and provide evidence on the nature of re-location decisions. We find that inventors were significantly more likely to have migrated from their state of birth than *both* high-skill and low-skill persons. We also document that conditional on moving, inventors tended to relocate to more densely-populated and financially-developed places that were more likely to foster innovation.

Our data show that the early exit of unproductive inventors led to positive selection, while inventors produced their highest quality inventions (measured by patent citations) early in their careers. Moreover, the probability of exit increased in late stages of the life cycle, as inventors faced obsolescence through creative destruction. This type of churning activity can play an important role in the growth development process, as shown by the empirical literature on productivity and firm dynamics (e.g., Haltiwanger (2012)) and theoretical models on firm entry and exit (e.g., Acemoglu et al. (2013), Jovanovic (1982) and Hopenhayn (1992)).

We also find that inventors had high incomes, even after controlling for their observable characteristics. Inventors had three times higher labor income on average and had a steeper earnings profile over their life cycle. Fully 59% of inventors were in the top decile of the overall income distribution. We identify strong returns to the quality of innovation: inventors with higher citation-adjusted patents received higher wage income. Furthermore, we find that wage income is related to the life cycle of invention. Young inventors with a longer career ahead of them were paid in anticipation of future productivity whereas for older inventors both future and past productivity predict wage income.

We find strong evidence to suggests that inventive activity translated into faster economic growth. We study the relationship between patented inventions and long-run growth across states over 100 years between 1900 and 2000. Our results show that the link has been strongly positive and economically sizable. Estimates suggest that if two states had the same initial GDP per capita in the beginning of the period and one state was at the 10th percentile and the other at the 90th percentile of the innovation distribution (Mississippi vs New Jersey, for instance), this could lead to 26% higher GDP per capita in the innovative state after 100 years. We attempt to establish causality by exploiting a major shift in innovation activity during World War II when the Office of Scientific Research and Development funded a program of unanticipated technological developments.

Finally, we study regional income dynamics as an outcome measure in relation to prior-period patenting activity. We focus on various measures of inequality: the 90/10 ratio, the Gini coefficient and the top-1% income share. We also construct a measure of social mobility using information in the 1940 Census that focuses on the fraction of those with a low-skill father who themselves have a high-skill occupation. We find that innovative regions in the U.S. had lower income inequality measured as the 90/10 ratio or the Gini coefficient, yet the top income share features a U-shaped relationship with state innovation. In general, the most innovative states had higher levels of social mobility.

Overall, our analysis uses novel historical microdata linked to regional aggregates to provide key micro and macro-level facts to inform critical questions in the study of technological progress and long-run economic growth. The remainder of the paper is organized as follows. Section 2 outlines our data. In Section 3, we present a simple theoretical framework to inform our empirical exploration. Section 4 presents our micro-level empirical results and Section 5 presents the macro findings. Section 6 concludes. Appendices A to D provide a detailed description of the data used, our matching methodology and additional robustness checks.

2 Data Construction

Patents, Technology Areas and Citations

Patents are a commonly used measure of innovation in the empirical literature on technological change. A patent entry shows the surname, first name, middle initial(s) where relevant, state, city, county, and country of the applicant when the patent was granted. For example, Figure 2 shows the famous USPTO patent 223,898 for an electric lamp granted to Thomas Edison in Menlo Park on January 27, 1880. As patents represent transferable property rights, they may be assigned to an individual or firm other than the inventor; if this is the case, the assignee is also recorded on the document.[2] Figure 1 (Panel B) shows that the share of American patents that were assigned in this manner grew over time.

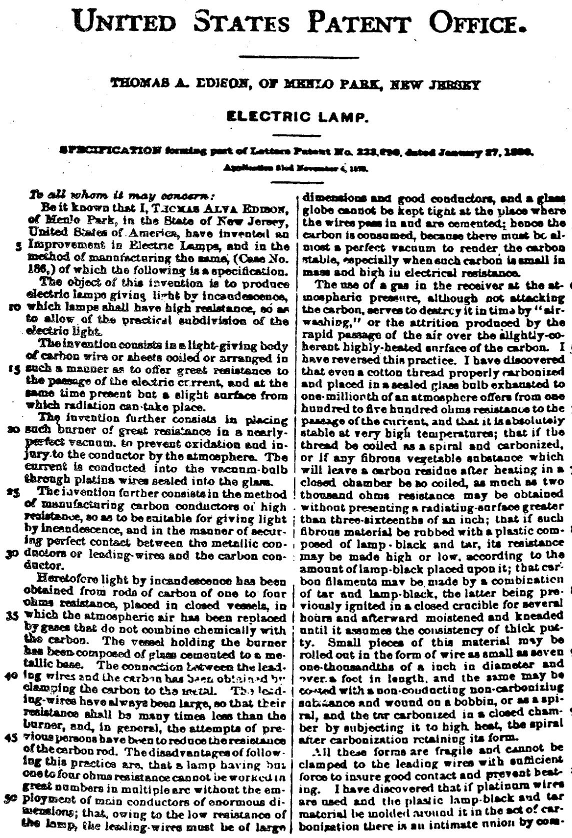
Using a combination of machine learning techniques and hand entry, we build a comprehensive collection of over 6 million U.S. patents granted between 1836 and 2004, which allows us to gain unique insights into the characteristics of U.S. inventive activity over long time horizons. The construction of these patent data are described in Appendix B.

In addition, we augment the information available from the original patent documents with two datasets. First, we use the USPTO's classification of patents to isolate the technology area of inventions. Whenever a new classification is introduced, existing patents are retroactively reclassified; therefore, this classification is consistent over time. Although patents may list several technological components, we only use the primary classes and subclasses for each invention. Second, we use historical patent citations to identify the most influential technological development. Our data include 3.7 million citations to patents granted between 1880 and 1940 from the population of patents granted between February 1947 (when front page citations began to be systematically recorded) and September 2008. Following Hall et al. (2001), we adjust citations to account for bias due to truncation or aggregate fluctuations in citation propensity.[3]

Several aspects of patenting are worth highlighting in the context of our link between inventors and the Census records. First, access to patenting was widespread. The cost of obtaining a U.S. patent was very low by international standards. Lerner (2002) estimates that to hold a patent

[2]Edison's patent was unassigned at the grant date.
[3]Details of the citation adjustment are provided in Appendix B.1.

FIGURE 2: PAGES OF USPTO PATENT NUMBER 223,898



UNITED STATES PATENT OFFICE.

THOMAS A. EDISON, OF MENLO PARK, NEW JERSEY

ELECTRIC LAMP.

SPECIFICATION forming part of Letters Patent No. 223,898, dated January 27, 1880.

Application filed November 4, 1879.

To all whom it may concern:

Be it known that I, THOMAS ALVA EDISON, of Menlo Park, in the State of New Jersey, United States of America, have invented an Improvement in Electric Lamps, and in the method of manufacturing the same, (Case No. 186,) of which the following is a specification.

The object of this invention is to produce electric lamps giving light by incandescence, which lamps shall have high resistance, so as to allow of the practical subdivision of the electric light.

The invention consists in a light-giving body of carbon wire or sheets coiled or arranged in such a manner as to offer great resistance to the passage of the electric current, and at the same time present but a slight surface from which radiation can take place.

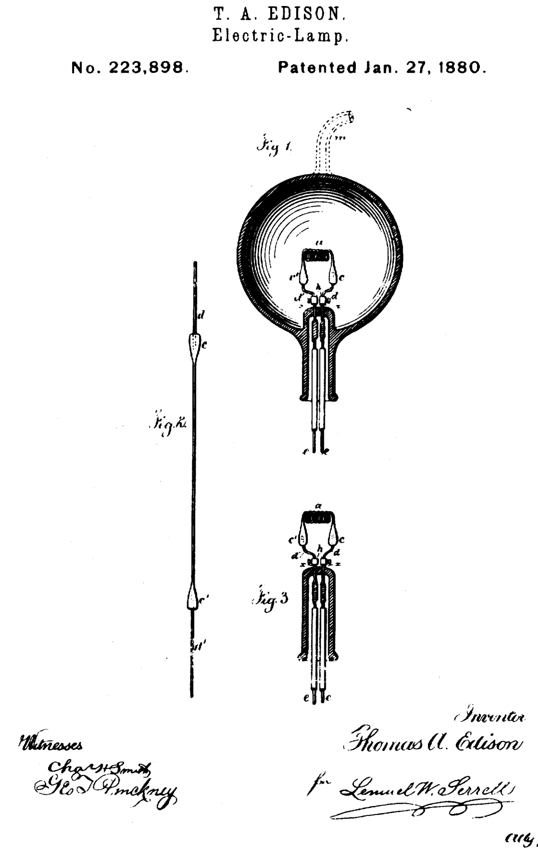
The invention further consists in placing such burner of great resistance in a nearly-perfect vacuum, to prevent oxidation and injury to the conductor by the atmosphere. The current is conducted into the vacuum-bulb through platina wires sealed into the glass.

The invention further consists in the method of manufacturing carbon conductors of high resistance, so as to be suitable for giving light by incandescence, and in the manner of securing perfect contact between the metallic conductors or leading-wires and the carbon conductor.

Heretofore light by incandescence has been obtained from rods of carbon of one to four ohms resistance, placed in closed vessels, in which the atmospheric air has been replaced by gases that do not combine chemically with the carbon. The vessel holding the burner has been composed of glass cemented to a metallic base. The connection between the leading wires and the carbon has been obtained by clamping the carbon to the metal. The leading-wires have always been large, so that their resistance shall be many times less than the burner, and, in general, the attempts of previous persons have been to reduce the resistance of the carbon rod. The disadvantages of following this practice are, that a lamp having but one to four ohms resistance cannot be worked in great numbers in multiple arc without the employment of main conductors of enormous dimensions; that, owing to the low resistance of the lamp, the leading-wires must be of large dimensions and good conductors, and a glass globe cannot be kept tight at the place where the wires pass in and are cemented; hence the carbon is consumed, because there must be almost a perfect vacuum to render the carbon stable, especially when such carbon is small in mass and high in electrical resistance.

The use of a gas in the receiver at the atmospheric pressure, although not attacking the carbon, serves to destroy it in time by "air-washing," or the attrition produced by the rapid passage of the air over the slightly-coherent highly-heated surface of the carbon. I have reversed this practice. I have discovered that even a cotton thread properly carbonized and placed in a sealed glass bulb exhausted to one-millionth of an atmosphere offers from one hundred to five hundred ohms resistance to the passage of the current, and that it is absolutely stable at very high temperatures; that if the thread be coiled as a spiral and carbonized, or if any fibrous vegetable substance which will leave a carbon residue after heating in a closed chamber be so coiled, as much as two thousand ohms resistance may be obtained without presenting a radiating-surface greater than three-sixteenths of an inch; that if such fibrous material be rubbed with a plastic composed of lamp-black and tar, its resistance may be made high or low, according to the amount of lamp-black placed upon it; that carbon filaments may be made by a combination of tar and lamp-black, the latter being previously ignited in a closed crucible for several hours and afterward moistened and kneaded until it assumes the consistency of thick putty. Small pieces of this material may be rolled out in the form of wire as small as seven one-thousandths of a inch in diameter and over a foot in length, and the same may be coated with a non-conducting non-carbonizing substance and wound on a bobbin, or as a spiral, and the tar carbonized in a closed chamber by subjecting it to high heat, the spiral after carbonization retaining its form.

All these forms are fragile and cannot be clamped to the leading wires with sufficient force to insure good contact and prevent heating. I have discovered that if platinum wires are used and the plastic lamp-black and tar material be molded around it in the act of carbonization there is an intimate union by com-



to a full term of 17 years in the United States in 1875 cost just 5% of the amount in the United Kingdom. Meanwhile, this cost in the U.S. was just 11% and 10% of the UK cost in 1900 and 1925, respectively. Moreover, inventors were actively encouraged to innovate and file for patents by the way the application process was configured. They could mail documents to the U.S. patent office in Washington, DC through the extensive network of post offices connecting the country (Khan (2009), Acemoglu et al. (2016b)), or use a large network of intermediaries (patent agents and lawyers) to administer the patenting process (Lamoreaux and Sokoloff, 1999).

Second, although patents could be sold in a market for technology that had flourished since the middle of the nineteenth century (e.g., Lamoreaux and Sokoloff (1999), Akcigit et al. (2016a)), the location of the original inventor is still recorded on the patent document. U.S. patent law stipulated that the "first and true inventor" be listed in the patent application even if the patent was assigned to another individual or firm at its grant date.

Third, the date of a patent application and the date of its eventual grant—when we observe the applicant's location—were quite close, at least for the early years of our study. In 1880 an average of 170 days elapsed between the filing and grant date; in 1900, 343 days elapsed. For this time period there is a reasonable alignment between the patent grant year and when an individual was observed in the Census year. By 1910, however, the average patent pendency period was almost a year-and-a-half (536 days). In 1930 it had extended to over 1,000 days and it was still over 800 days in 1940.[4] We would therefore expect to see more measurement error in our matching *ceteris paribus* for later years. Schmookler (1966) reports that it took about one-and-a-half years for an invention by an independent inventor to be produced.

Census Data

The release of the complete-count Census data by the Minnesota Population Center (IPUMS) provides an opportunity to examine a number of questions related to the historical development of innovation in the United States. We use the decennial Censuses in 1880, 1900, 1910, 1920, 1930, and 1940.[5] Our patent-Census matching exercise begins in 1880 because that is the first year a reasonable number of patent observations become available. Around 11,400 patents were granted to inventors residing in the U.S. in that particular year. The Census Bureau's 72 year lag release rule implies that the latest available Census is from 1940.

We view these data analogously to modern studies using administrative records such as Bell et al. (2015) who uncover major new facts about the nature of U.S. innovation. Not only is it possible to link the historical Census data with patent records (as we show below) but data on the entire population permits analysis of inventor life cycles relative to other sub-groups of individuals—for instance with different occupational skill levels. This type of systematic information across large groups of individuals for the entire United States has never before been available for long historical horizons. Nevertheless, although the Censuses present an especially useful source of data, it is also worthwhile to understand its potential limitations.

First, the quality of the Census records varies over time. While the Census included quality control procedures in an effort to ensure consistent enumeration, much depended on the way the Censuses were generally administered. For example, the 1920 Census was conducted in the winter (January 1st 1920) whereas the 1910 Census had been conducted in the spring (April 15th). Winter enumeration had a large effect on seasonal occupations like agricultural labor and movement to cities. Although we show in Appendix C that our match rate is lower for 1920 than for other years, the level of underenumeration is not sufficient to bias our results. Dorn (1937) estimates underenumeration in the native white population (which would be most relevant to the inventors in our dataset) of between 1% and 1.1%.

Second, beyond standard variables like the name and location of individuals, the information contained in the Censuses varies widely over our period of interest. Although a number of variables are commonly recorded across Census years such as age, race, gender and marital status, other variables are recorded in one year only to be dropped in another. As an example, occupation is listed in 1880 but not in 1900 or 1910. Generally, a wider array of variables are available in later years. Beginning in 1920, for instance, enumerators asked specifically about education (school attendance) and home ownership.

[4]By comparison, the average difference between a patent application and grant date for U.S. patents granted between 2008 and 2015 was 1,278 days.

[5]For a full description of the Census datasets, the variables they contain, and our attempts to clean them, see Appendix A. The 1890 Census was largely destroyed in a fire in 1921, while others records from this Census were destroyed under the intransigent Federal record management polices in place at the time. Only a limited set of 1890 Census schedules survived.

Third, some variables are subject to measurement error. An advantage of the 1940 Census is that it questioned individuals about income. Prior to the availability of these data, researchers routinely imputed incomes by assigning individuals the median income in their reported occupational category (Abramitzky et al., 2014). But, the income data must be carefully interpreted. Enumerators were instructed to report annual incomes of greater than \$5,000 at \$5,000+. Hence, the data are "top-coded." For example, Melvin De Groote, the superstar inventor profiled above, reports this level of income in the 1940 Census (average income was \$1,368). Furthermore, Petro (2016) finds that "if a farmer worked for himself and sold his crops, he did not report that money" in the 1940 Census. By the same token we assume that inventors selling their inventions would not have reported this as income.

Approach to Matching Patent and Census Records

The main challenge associated with matching inventors on patent documents to individuals listed in the Census is the absence of a unique identifier across datasets. The first Social Security number was issued in 1936. Although a supplementary question was asked in the 1940 Census about whether a person had a Social Security number, the number itself was not recorded. Furthermore, Social Security numbers were not included in patent documents at the time.

In both the patent and Census datasets we observe variables denoting surname, first name, initial, state, city and county. This vector of information provides a basis for our matching. Of course, the challenge of matching observations without unique identifiers is self-evident, but we can still limit the likelihood of matching "false positives" by restricting our analysis to only those observations where we match precisely across a range of our matching variables. We proceed in two steps. First, we adopt a "basic" matching approach where the criterion for matching is that the inventor listed on the patent has the same first name and surname as the individual in the Census, and lives in the same state. Naturally, this leads to repeated individuals in some cases.

Therefore, we next adopt a "refined" matching approach. In addition to the criterion in our basic match we require additionally that individuals listed on the patent document and individuals in the Census reside in the same county. Then, if there are still many observations for a given inventor, we first check if there is an inventor which has the same middle initial in both the patent and Census datasets, and we keep that inventor if there is a match. We then keep only Census inventors who live in the same city or township as is listed on the patent document, if one exists. Next, we ask if there is any matched inventor between 16 and 85 years old. If so, we keep that inventor only. Finally, we repeat the age refinement, keeping only matched inventors between 18 and 65 years old, if one exists. In other words, to be in our final dataset requires that individuals match systematically on surname, first name, (where relevant initial), state, and county. Although there are still data matching issues we cannot overcome—for example, sometimes the Census uses registration areas (e.g., Precincts or Districts) rather than cities, making it impossible for us to identify the right individual by location—our matching rates are encouraging overall. We match an average of 46 percent of patentees in the Census with a high of 62 percent in 1880 and a low of 34 percent in 1920. A detailed description of the matching process is provided in Appendix C.

2.1 Summary Statistics

As a precursor to the main analysis we present descriptive statistics on our data. In keeping with our approach of examining U.S. innovation from micro and macro perspectives, we structure these data to characterize inventiveness at both the individual inventor and state levels.[6]

Micro-level Summary Statistics

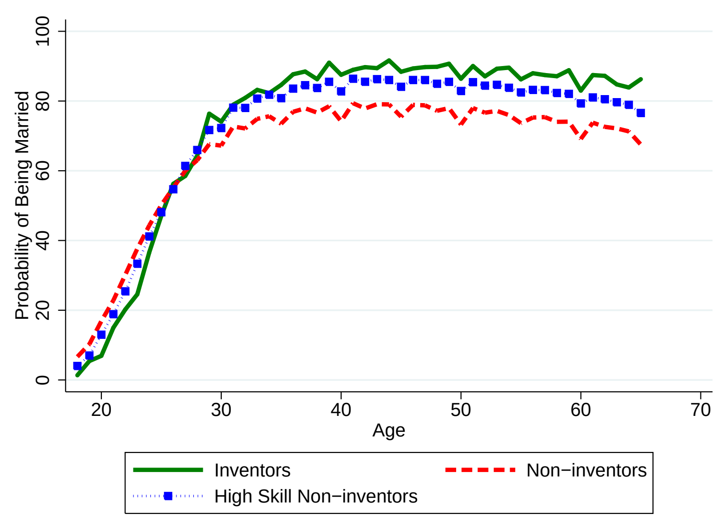
Table 1 shows inventors were more likely to be white males. During our time period women's involvement in the labor market was generally restricted to positions like office and clerical work (Goldin (2006)). Khan and Sokoloff (2004) found only one female inventor in their list of 400 superstar U.S. inventors who were born before 1886. Cook (2011) finds that while African American inventors often made important technological discoveries during the nineteenth and early twentieth centuries, they were much less likely to do so in closed environments such as places that implemented segregation laws.

Table 1: The Characteristics of Inventors

	Inventors	Full U.S.
Percent White	97.9%	89.4%
Percent Black	1.8%	9.1%
Percent Male	97.9%	51.0%
Single	16.1%	27.7%
Married	80.2%	65.4%
Percent 19-25	8.4%	22.6%
Percent 26-35	23.8%	27.5%
Percent 36-45	31.0%	22.5%
Percent 46-55	24.1%	16.6%
Percent 56-65	12.7%	10.8%
Av. # Children: $\leq$ 35 yrs old	1.9	2.3
Av. # Children: $>$ 35 yrs old	3.2	4.7
Percent Interstate Migrant	58.8%	42.8%
Percent International Migrant	21.1%	17.4%
Percent Of Population	0.02%	99.98%

Notes: We use all matched census records to construct this table. Age, race, marital status, and migrant status are reported for all years. Fertility is reported only in 1910 and 1940. Source: 1880 through 1940 Historical Census Data, USPTO patent records.

Figure 3: Family Decisions: Probability of Being Married



Notes: This figure plots the probability that an individual is married over their life cycle using data averaged across our six census years. Source: 1880-1940 Historical Census Data, USPTO patent records.

Inventors tended to be married, middle-aged, and had fewer children early in their careers. Theoretical models specify that commitment to a spouse soaks up time and effort, and that if married partners did not gain from a union then they would remain single (e.g., Becker (1974)). To the extent that inventors' life cycle dynamic created tradeoffs with respect to time allocation, one might expect inventors to delay fertility and marriage.

Anecdotally, some of the most prolific inventors were against marriage. Nikola Tesla commented in the *New York Herald* in 1897 "I do not believe an inventor should marry, because he

[6]Due to space constraints, we report a limited amount of summary statistics here. More details can be found in the NBER Working Paper Version of our paper Akcigit et al. (2017b).

has so intense a nature, with so much in it of wild, passionate quality, that in giving himself to a woman he might love, he would give everything, and so take everything from his chosen field." Tesla went on to argue that "I do not think you can name many great inventions that have been made by married men." However, other great inventors did marry. Elias Howe (1819-1867), the inventor of the sewing machine, married when he was 21 years of age. Thomas Edison married first at age 24 and within a year had developed the revolutionary quadruplex telegraph for sending multiple messages simultaneously over a single wire. Following the death of his first wife, Edison married again at age 39.

Figure 3 shows that inventors delayed marriage relative to the population as a whole, although inventors did marry (or stay married) at a higher rate than non-inventors at older ages. Figure 3 also plots the probability of marriage for those working in a high-skill occupation, such as doctors and lawyers. The figure shows that inventors' marriage decision mirrors that of this group almost one-for-one. This comparison suggests that inventors' difference from the rest of the population along this dimension is driven by underlying skill differences and human capital investment choices. This similarity in observable marriage patterns with high-skill workers can be reconciled with theoretical models of marriage markets like Bergstrom and Bagnoli (1993), where high-wage men gain by delaying marriage relative to low-wage men because accumulated income is a signal of quality when searching for the best partner.

Table 1 indicates that inventors delayed fertility relative to the average American: 72.9% of inventors had a child before the age of 35, while 80% of non-inventors had a child by this time. Of course the relationship between delayed marriage and fewer children is mechanical. As Becker (1974, p.22) points out, "the age of entry [into marriage] would be earlier the larger the number of children desired."

Table 1 also shows higher levels of interstate and international migration among inventors. Migrants can bring new ideas, expertise, and specialized labor to an area, all of which can facilitate the production of patented innovations. For example, Moser et al. (2014) estimate that German emigres who fled the Nazi regime provided a significant boost to U.S. invention during the twentieth century. In a companion study (Akcigit et al., 2017a) we explore the role of international migrants, finding that technology areas with higher levels of foreign-born expertise grew much faster between 1940 and 2000 than otherwise comparable technology areas.

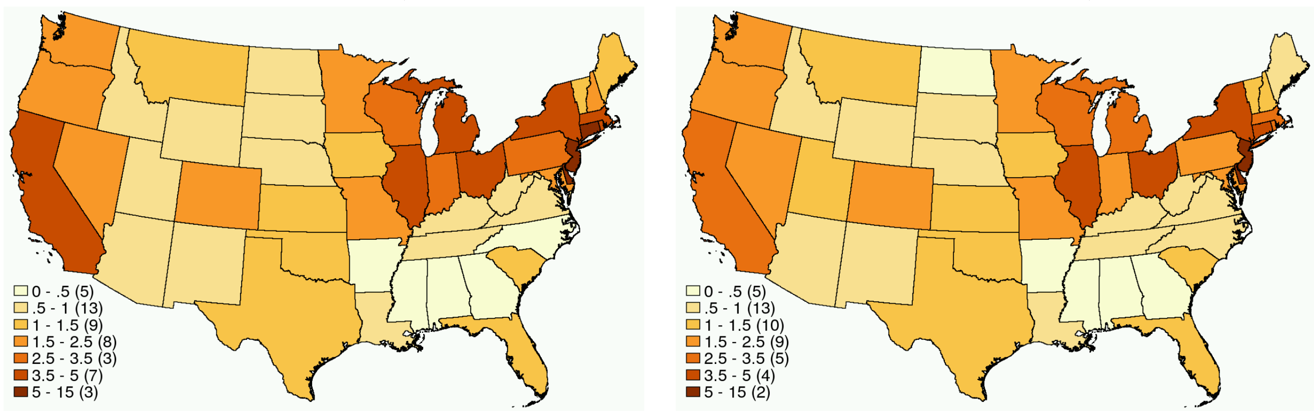
Finally, the base of Table 1 shows that inventors represented a very small share of the population—just 0.02%. By comparison, in all years for which occupation data is reported in the Census, 0.46% of the working age population was a doctor or a lawyer. When knowledge diffuses rapidly, inventors developing breakthrough inventions can have a large influence on economic growth. During the U.S. golden age and major epochs of economic development more generally, the technological ingenuity and innovative capabilities of the minority tended to matter the most (e.g., Squicciarini and Voigtlaender (2015)).

Macro-level Summary Statistics

At the state-level, Figures 4A and 4B illustrate the geography of inventiveness defined as patents and inventors per 10,000 people in 1940. Both figures reveal concentrations of activity in rust-belt manufacturing areas, which mirrors the distribution of industrial activity at the time (Glaeser, 2011). California also stands out as a center of innovation for most of the years we observe. This is not caused by sparse population counts mechanically inflating the patent and inventor counts. While Los Angeles ranked as the 36th largest city in the U.S. in 1900, it was ranked number 10 in 1920 and number 5 in 1940. Figures A-7 and A-8 in Appendix D show that the geographic patterns displayed in Figure 4 are remarkably stable across our six census years.

Figure 4: The Geography of Inventiveness

Panel A: Patents per 10,000 People | Panel B: Inventors per 10,000 People



Notes: Figure maps the number of patents (panel A) or inventors (panel B) per 10,000 residents in each state of the mainland U.S. in 1940. Darker colors represent more inventive activity per resident. Patent data come from the USPTO's historical patent files, while population counts are calculated using the U.S. Census. Appendix D reports similar maps in different decennial census years.

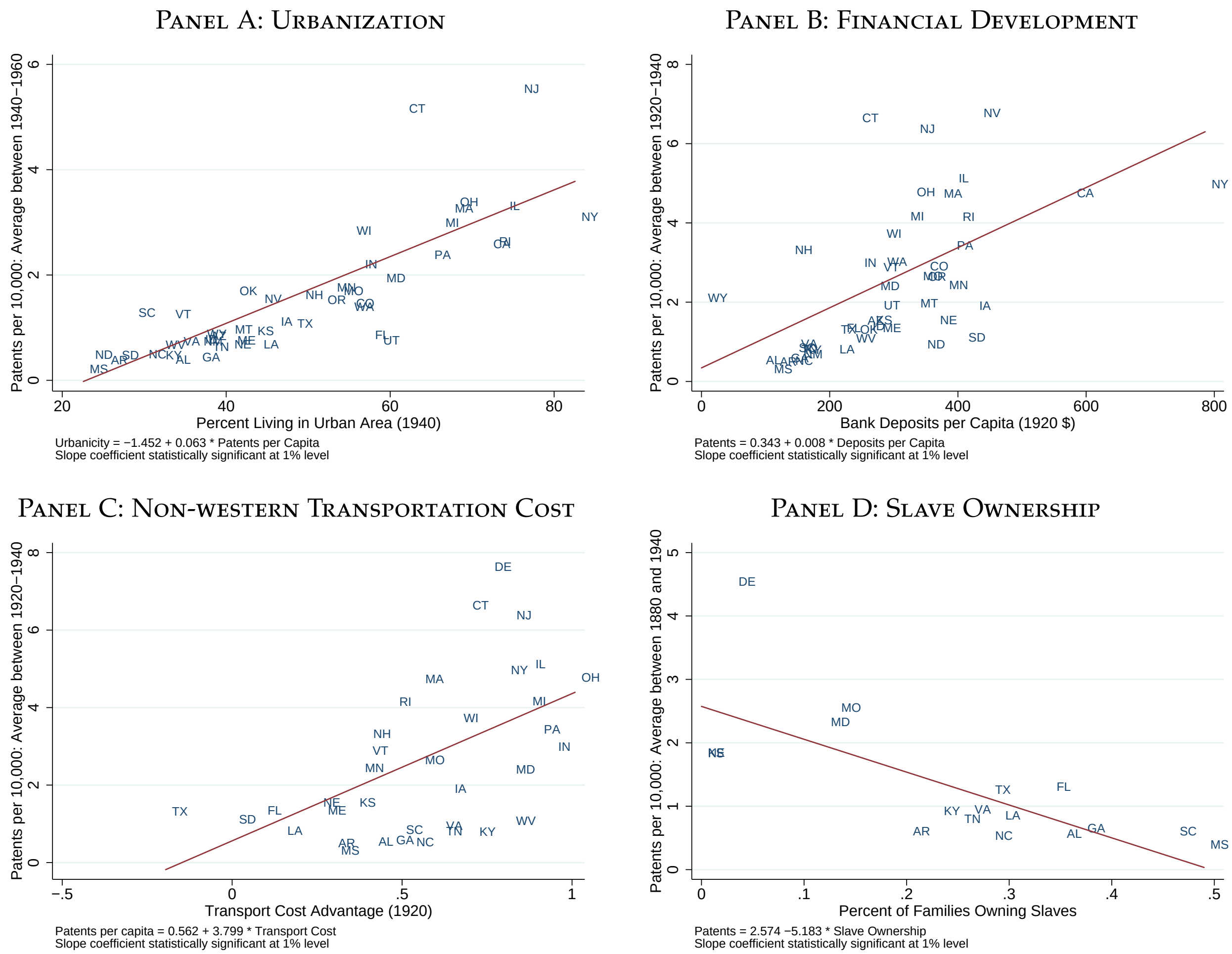
Some states, by virtue of favorable resource endowments, institutions or culture may have been particularly conducive to inventive activity overall. Figure 5 highlights the importance of commonly postulated drivers of innovation: population density, financial development, geographic connectivity and social structure measured by association with slavery.

Panel A shows population density was much higher in the most inventive states.[7] This finding is consistent with human interaction being key for human capital accumulation and economic growth (e.g., Lucas (2009), Alvarez et al. (2013), Lucas and Moll (2014), Perla and Tonetti (2014), Akcigit et al. (2016b)). The agglomeration literature asserts that physical proximity can promote creativity and the exchange of ideas among inventors (Carlino and Kerr (2015)).

Panel B shows a strong positive correlation between financial development and innovation. Although private transactions between investors and inventors are not observable systematically, and most later stage R&D is financed by firms internally, we can measure the general health of a state's financial sector using Federal Deposit Insurance Corporation data on bank deposits per capita in 1920. We find that a healthier financial sector is correlated with greater innovation levels. Higher levels of financial development are typically associated with faster rates of economic growth (e.g., King and Levine (1993)).

[7] Because the Census adopts a low threshold for urbanization as places that encompass at least 2,500 people, we repeated the analysis at different thresholds with similar results. When we use a threshold of 5,000 people the slope coefficient is 0.061 and at 10,000 people the slope coefficient is 0.056. All are significant at the 1% level.

Figure 5: State Characteristics and Innovation



Notes: Panel A shows the relationship between the percent of individuals living in an urban area in the 1940 census and average patents per capita between 1940 and 1960. Panel B shows the relationship between the amount of deposits per capita in 1920 dollars and average patents per capita between 1920 and 1940. Banking data originate from the FDIC dataset, downloaded from the University of Michigan's ICPSR repository (number 0007). We use 1920 data to remove the influence of the Great Depression from our data. In both panels Delaware excluded as an outlier to for visibility. Panel C plots the relationship between outgoing shipment costs and innovation. The horizontal axis measures the number of standard deviations below the mean cost of transporting one ton of goods to other states. As Western states were less integrated in the national economy at the time, the figure only plots the relationship for states with average outbound transport cost under $18 per ton. Panel D plots the relationship between the percent of families which owned slaves in the 1860 census, and average patents per capita between 1880 and 1940. Sources: U.S. Census, FDIC, Donaldson and Hornbeck (2016), USPTO patent records.

Another important dimension for innovation is access to other geographical regions. This could increase both the market size for innovation and the flow of knowledge spillovers. Donaldson and Hornbeck (2016) measure the increased level of market access caused by an expansion of the U.S. railroad network while Perlman (2016) finds strong effects on invention and agglomeration from the nineteenth century development of railroads. In Panel C we find a strong relationship between a state's geographic connectivity and its level of patenting in the older, more integrated East Coast and Midwestern states where economic activity was concentrated.

Cultural differences may be an important determinant of a region's innovative activity and growth. An especially important aspect of openness of a society to innovation and economic growth could be seen from its approach towards slavery. Panel D highlights that states with high slave populations in 1860 were disproportionately among the least inventive in the United States between 1880 and 1940. A lack of cultural freedom to deviate from established norms can strongly inhibit innovation. Innovative places tend to be more open to unconventional and disruptive technological ideas (e.g., Florida (2002), Acemoglu et al. (2014)).

3 A Motivating Theoretical Framework

These data provide us with multiple directions for novel empirical investigation. To frame our analysis, we present a simple model of innovation that integrates insights from standard theories in the existing endogenous growth literature. Our aim is not to develop a new model, but rather to use existing frameworks in the growth literature to organize our empirical analysis.

Basic Environment. Time is continuous. At any instant t, the final good Y_t is produced according to the following Cobb-Douglas production function in a perfectly competitive market

$$Y_t = \exp\left(\int_0^1 \ln y_{jt} dj\right) \tag{1}$$

where y_{jt} denotes the quantity of variety j at time t. We normalize the price of the final good to 1 at every instant t without any loss of generality. Each variety is produced by a monopolist who owns the best technology in variety j according to the following production function

$$y_{jt} = q_{jt} l_{jt}.$$

In this expression, l_{jt} is the amount of production workers hired by monopolist j at time t and q_{jt} is labor productivity. This production function implies that the marginal cost of producing one unit of y_{jt} is

$$MC_{jt} = w_t / q_{jt},$$

where w_t is the wage rate paid to each production worker.

Innovation improves labor productivity by moving firms up a quality ladder, in a manner similar to the model of Aghion and Howitt (1992) and Grossman and Helpman (1991). More specifically, a new innovation in a particular variety j increases labor productivity of workers in that variety by a multiplicative factor $(1 + \lambda)$. Therefore, during a small time interval Δt, quality improves according to the following law of motion

$$q_{jt+\Delta t} = \begin{cases} (1+\lambda)\, q_{jt} \text{ if there is a successful innovation,} \\ q_{jt} \text{ otherwise.} \end{cases}$$

When there is a new innovation in j, the latest inventor and the previous incumbent enter into Bertrand price competition.

Profits. Now we solve for equilibrium profits, which in turn determine the incentives for innovation. Note that the Cobb-Douglas final good production function (1) generates the standard unit-elastic demand for each variety, as $y_{jt} = Y_t / p_{jt}$. Given this demand, Bertrand competition implies that the market leader chooses to set the price of variety j to be the marginal cost of the previous incumbent. Since innovation increases labor productivity by a constant multiplicative factor $(1+\lambda)$, the previous incumbent's productivity is given by $q_{jt}/\left(1+\lambda\right)$. Therefore, we may express the price of variety j as $p_{jt} = \left(1+\lambda\right) w_t / q_{jt}$.

Imposing this pricing strategy yields the optimized profit for an incumbent firm

$$\pi_{jt} = \left[p_{jt} - MC_{jt}\right] y_{jt} = \frac{\lambda}{1+\lambda} \times Y_t. \tag{2}$$

This expression implies that profits will be independent of the variety index j. Note that Y_t increases the demand for each variety and hence leads to higher monopoly profits through an increased market size effect. We will return to this feature below.

Substituting the monopoly quantity into (1), we find that the wage rate is simply

$$w_t = \frac{Q_t}{1+\lambda} \tag{3}$$

where $Q_t \equiv \exp\left(\int_0^1 \ln q_{jt} dj\right)$ is the aggregate quality index of this economy. This implies that the equilibrium level of output is simply

$$Y_t = Q_t \times L_P \tag{4}$$

where L_P is the measure of production workers in the economy.

Innovation. We now turn to the dynamics of the model. There is a mass $1+L$ of individuals in this economy. A measure 1 of these individuals operate as business owners as described above. The remaining measure L may choose to be a production worker $i \in L_P$ or innovator $i \in L_I$ such that $L_I \cup L_P = L$ and $L_I \cap L_P = \varnothing$. Production workers earn the wage rate w_t. If individual i decides to become an inventor, he/she pays a cost $c_i Q_t$ and receives a patentable innovative idea with probability η. For tractability, we assume that the cost is known to the agent before they make their career choice, and is uniformly distributed between 0 and $\beta : c_i \sim U\left[0, \beta\right]$. Note that every person in the economy is more likely to draw a high cost when β increases, therefore β proxies for macro-level factors that affect innovation.

Let us denote the value of a successful innovation by V_t. In this case, a person decides to become an inventor if and only if the expected value of becoming an inventor is greater than the outside option of being a production worker:

$$\eta V_t - c_i Q_t > w_t.$$

This implies that there is a cut-off $c^* = \left(\eta V_t - w_t\right)/Q_t$ below which individuals decide to become inventors:

$$i = \begin{cases} \text{inventor if } c_i < c^*, \\ \text{production worker otherwise.} \end{cases} \tag{5}$$

Using the equilibrium wage rate (3) , this cut-off is simply

$$c^* = \eta \frac{V_t}{Q_t} - \frac{1}{1+\lambda}. \tag{6}$$

Value of Production. Let ρ be the common rate of time preference of the household that has logarithmic utility over consumption in the economy, and denote the flow of new patents by $\tau \equiv \eta L_I = \eta L \frac{c^*}{\beta}$. Since every new patent displaces an incumbent, τ is also the rate at which incumbents are replaced by new entrants. Therefore incumbents discount future payoffs at the rate $\rho + \tau$, so that the value of holding a patent as an incumbent is[8]

$$V_t = \frac{\pi_t}{\rho + \tau}. \tag{7}$$

Let the share of the population that works as inventors be given by $s \equiv L_I/L$. We may combine (2) , (4) , (6) , and (7) to express implicitly the equilibrium fraction of inventors s in the society:

$$\underbrace{\frac{(1-s)\,\lambda}{1+(1+\lambda)\,\beta s} - s}_{\equiv F(s,\beta,\lambda)} = \frac{\rho}{\eta L} \tag{8}$$

Note that the left-hand side of this equation, which we define as $F(s, \beta, \lambda)$, is increasing in the innovation step size λ, and decreasing in both the share of inventors in the population s, and the maximum of the inventor cost distribution β. Finally, it is straightforward to show[9] that the growth rate of aggregate output ($g = \dot{Y}/Y$) is increasing in the number of patents produced

$$g = \tau \times \ln(1+\lambda). \tag{9}$$

Discussion of the Model

In this model, η increases the likelihood of innovation, so that those with a higher probability of successful innovation will be more likely to innovate. Empirically, the highest-skilled individuals in an economy, possessing an intimate knowledge of their field, may therefore be most likely to innovate. This suggests a strong relationship between education and innovation, as confirmed anecdotally by the example of Melvin De Groote (see also Fact 1):

$$\frac{dPr(Being\ an\ inventor)}{dEducation} > 0.$$

In addition, expression (5) shows that there are micro-level factors c_i that determine the

[8] More formally, the value function (Hamilton-Jacobi-Bellman equation) of being an incumbent is

$$rV_t - \dot{V}_t = \pi_t - \tau V_t.$$

Imposing the fact that in steady state we have $\dot{V}_t = gV_t$, and that the household has logarithmic utility first delivers $\rho = r - g$ and then equation (7).

[9] Note from equation (4) that $g = \dot{Q}_t/Q_t$. Moreover $\ln Q_t = \int_0^1 \ln q_{jt} dj$. Since every line receives a new innovation at the rate τ, after a small time interval Δt we get $\ln Q_{t+\Delta t} = \ln Q_t + \tau \Delta t \ln(1+\lambda)$. Subtracting $\ln Q_t$ from both sides, dividing by Δt and taking the limit as $\Delta t \to 0$ delivers the result.

probability of becoming an inventor. If a person faces a higher idiosyncratic cost, he/she would be less likely to become an innovator. These individual-level costs may be lower for those with additional institutional knowledge of the innovation process, such as those with a father as an innovator. Likewise, favorable parental income could lower the cost of financing or introduce a child to the social networks necessary to market a new innovation (see Fact 2). We can summarize this theoretical implication as:

$$\frac{dPr(Being\ an\ inventor)}{dMicro\ costs} < 0.$$

Relatedly, the innovativeness of an economy is also falling in macro-level cost parameter β, which governs the distribution of the fundamental costs inventors face. Specifically, the model predicts that $\frac{d\tau}{d\beta} = \eta L \frac{ds}{d\beta} < 0$. This result in intuitive. Increases in β imply that individuals in the economy face a higher cost of innovation on average. Empirically, such costs might arise due to a lack of sufficient financing, or from social stigma and an unwillingness to accept disruptive ideas (recall Figure 5 and see also Fact 3). Therefore,

$$\frac{dPatents}{dMacro\ costs} < 0.$$

Additionally, in this standard model innovation incentives decline strongly once an entrant becomes an incumbent. This is also known as the *Arrow's replacement effect* in the endogenous growth literature. We study this empirically, by considering how innovation quality changes as an inventor becomes well-established (see Fact 5):

$$Innovativeness\ of\ new\ inventors > Innovativeness\ of\ incumbents.$$

A number of additional interesting predictions of this model can also be studied empirically. For instance, the return to innovation π_{jt} increases in the quality of innovation λ (see equation (2)). This implies that inventors' compensation should increase in patent quality (see Fact 6 for empirical evidence):

$$\frac{dInventor's\ compensation}{dPatent\ quality} > 0.$$

In this model, better technologies are introduced through new patented innovations which makes production workers more productive. Therefore equation (9) predicts that patenting should be strongly correlated with aggregate growth. Motivated by this observation, our empirical analysis (e.g., Fact 7) examines the link between patents (τ) and economic growth (g):

$$\frac{dEconomic\ growth}{dPatents} > 0$$

Finally, our model also implies that in the absence of innovation, every person holds the same position in the economy, i.e., business owners will remain business owners and assembly line workers will remain assembly line workers. However, innovation generates turn-over in society. Therefore our model implies that more patenting should relate to more social mobility in society (see Fact 9):

$$\frac{d\text{Social mobility}}{d\text{Patents}} > 0.$$

In sum, the model acts as a guide to our empirical work by underscoring key micro and macro-level relationships that are important to explore in the data.

4 Inventors' Background, Lifecycle and Earnings

We now use our dataset of patents matched to the Censuses to explore factors influencing the probability of becoming an inventor. We examine the personal background of inventors, paying special attention to their educational attainment and age, migration decisions, and entry and exit over the cycle of their inventive career.

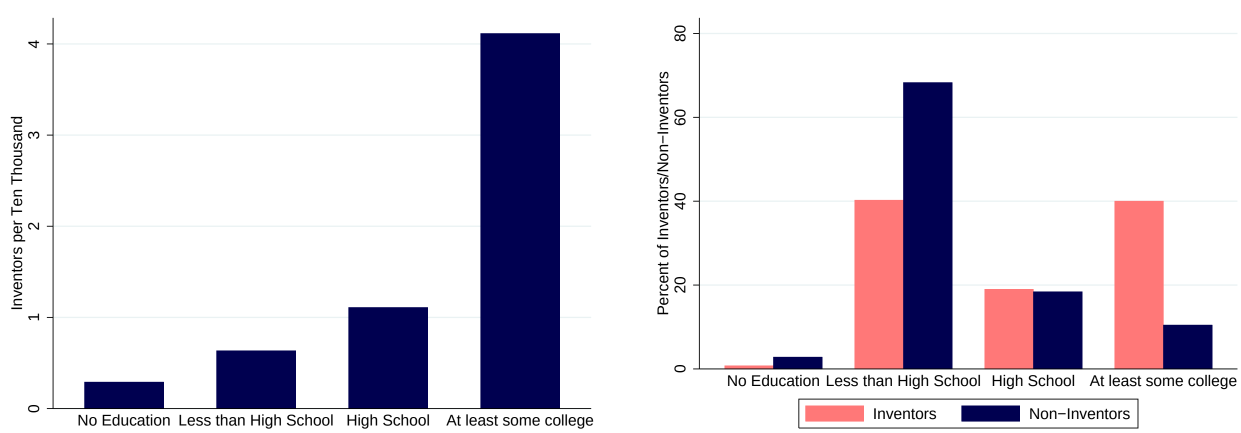
Fact 1 ***Inventors were more educated on average and were most productive between the age of 36 and 55.***

One of the main channels through which education affects economic growth may be its impact on innovation. Figure 6 shows the number of inventors per 10,000 people within each education group.

Figure 6: Education and Probability of Becoming an Inventor

Panel A: Inventors per 10,000 by Education

Panel B: Percent of Inventors in Each Education Category

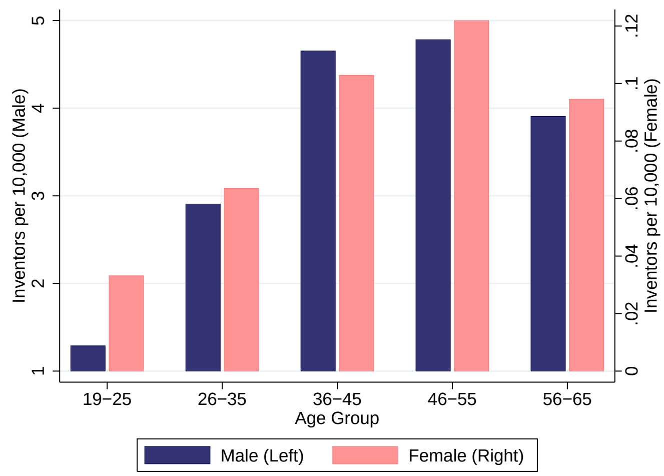


Notes: Figure plots the education of inventors and non-inventors in the 1940 census, the only census in our sample to provide sufficiently granular education information. Panel A plots the inventors per 10,000 people by education category. Panel B plots the percent of inventors and non-inventors that fall into each educational category. Source: 1940 Historical Census Data, USPTO patent records.

While education seems to be an important determinant of becoming an inventor, the effect is particularly strong at the college degree level. Although the 1940 Census tended to overstate education levels (Goldin, 1998) the differences we see between categories are large. For example, an individual with at least a college degree is four times as likely to become an inventor than an individual with just a high-school diploma. Indeed, 40% of inventors had a college degree in 1940, compared with just 10% of the non-inventor population.

The opportunity cost of education is time spent in an active career. In theoretical models of education, individuals face a tradeoff between the benefits of higher education, which accumulate over the life cycle, and the costs which are incurred early on (Becker (1967)). By extension economic growth can be affected by the tradeoff inventors face between acquiring human capital to innovate and the potential delays this creates in the production of new technological discoveries that, in turn, benefit society. Jones (2010) argues that if true breakthroughs are developed by younger cohorts of individuals, the growth-slowing delay effect can be pronounced, especially if more human capital is required for the production of creative ideas as the demands of developing novel innovations increases over time. He finds that the age of great invention shifted upwards by about half a decade over the twentieth century.

Figure 7: Probability of Innovation by Age



Notes: Figure shows the average life cycle of inventiveness over the years 1880 to 1940. It plots the number of inventors per 10,000 individuals by gender. The dark blue bars plot the number of male inventors per 10,000 males against the left axis, while the bright red bars plot the number of female inventors per 10,000 females against the right axis. Source: 1880-1940 Historical Census Data, USPTO patent records.

We find that inventors were most productive between ages 35-55 as illustrated in Figure 7. This is true for males and females, although female inventors were rare at this time. Interestingly, as shown by Sarada et al. (2016) the average age of invention in 1900 was approximately 40 years old, about what it is today.

Fact 2 ***Father's income and father's education were highly correlated with becoming an inventor, especially through the effect on the level of a child's education.***

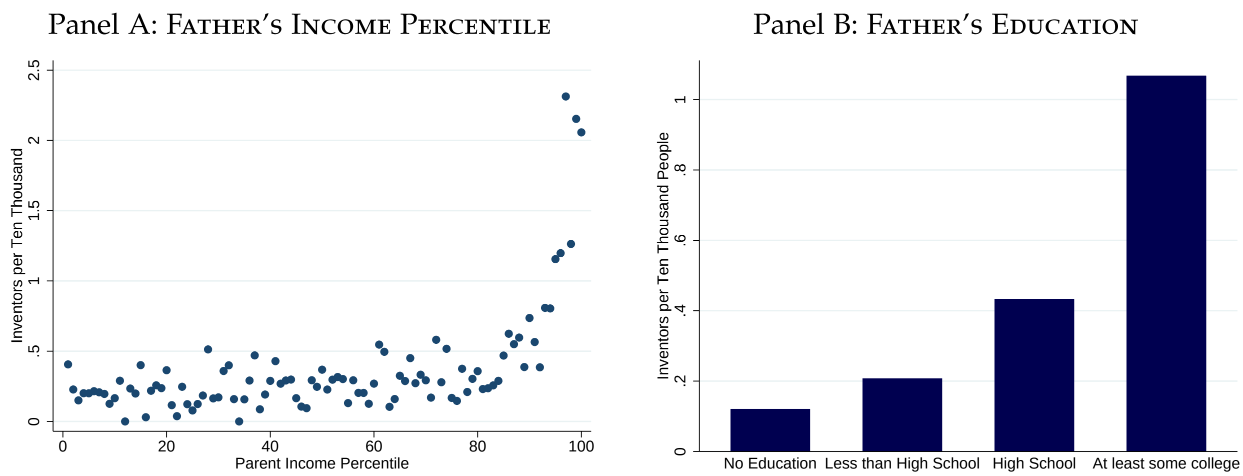
We now examine family backgrounds. Does parental affluence matter for the propensity to become an inventor? If so, through what mechanisms might this operate? In this section we rely heavily on our parent-child matched dataset.[10] Because this covers individuals residing in the same household, we are capturing inventors early in their career. Home-leaving ages increased noticeably during the early twentieth century only starting to decline after World War II. Using Census data Gutmann et al. (2002) find that in 1940 the median home-leaving age for white males

[10]For details of its construction, see appendix A.2.

was 24 whereas 85% of unmarried white males lived at home between ages 15 and 29.[11]

Figure 8 illustrates the relationship between parental affluence and the propensity to become an inventor. Panel A shows a strong association between the probability of becoming an inventor and father's income, especially for the highest-income fathers. Because this figure illustrates the contemporaneous relationship between the income of the father when the inventor had already entered their career, Panel B reproduces the result of Panel A using father's education to proxy for income when the inventor was a school-age child. The convex relationship between parental income (measured directly or using our proxy measure) and the propensity to become an inventor is striking in its ubiquity. Aghion et al. (2015b) and Bell et al. (2015) document remarkably similar patterns in modern data from Finland and the United States, respectively. The persistence of this relationship across time periods, geographies, and institutions is among the most noteworthy facts in this new literature on the backgrounds of inventors.

Figure 8: Parental Affluence and the Probability of Becoming an Inventor



Notes: Figure plots the number of inventors per 10,000 people by their father's percentile of wage income in the 1940 census (Panel A) or their father's education level (Panel B). Only individuals successfully matched to their fathers are included in this plot. Wage income percentiles are calculated using the full sample of matched fathers in the U.S. Source: 1940 Historical Census Data, USPTO patent records.

Several mechanisms can plausibly drive the patterns illustrated in Figure 8. If education was an important determinant of innovation, then the fact that only wealthy individuals had access to education could imply that credit constraints were binding for low-income families (e.g., Celik (2015)). Furthermore, credit constraints may inhibit the ability of prospective inventors to raise starting capital to develop their ideas. Alternatively, it is possible that high income parents interact in better-connected social circles, permitting their children to access high-quality funding, labor, and marketing resources. Finally, high income parents may have useful skills, knowledge, or genes which they pass on to their children.

We provide insight into some of these potential mechanisms through Table 2, which examines the relationship between fathers and sons using linear probability regressions. The dependent

[11] The focus on individuals who still live with their parents makes our study of family dynamics comparable to that of Bell et al. (2015), who limit themselves to a study of young inventors aged 28-32 in 2012.

variable is an indicator for being granted at least one patent, scaled by a factor of 100 for legibility, and in each regression, we include controls for race, sex, migration status, a quadratic in age, state fixed effects and father's age. Column 1 establishes a strong positive correlation between the father being an inventor and the child being an inventor. Column 2 adds contemporaneous parental income as measured in the 1940 Census.[12] Those with an inventor father are 0.16 percentage points more likely to become an inventor than those without an inventor father. Similarly, those with a father in the top 5 percent of the income distribution are 0.008 percentage points more likely to become an inventor. Given that just 0.02% of all individuals were inventors, these constitute large effects.

Table 2: Who Became an Inventor?

	(1)	(2)	(3)	(4)	(5)
Father Inventor	0.161**	0.159**	0.154**	0.159**	0.155**
	(0.075)	(0.076)	(0.076)	(0.075)	(0.075)
Father Income $90^{th}-95^{th}$ %ile		0.003**	-0.001		
		(0.001)	(0.001)		
Father Income 95^{th} %ile and above		0.008***	-0.000		
		(0.002)	(0.002)		
Father: High School Graduate				0.005***	-0.002
				(0.001)	(0.001)
Father: At least Some College				0.009***	-0.002**
				(0.002)	(0.001)
Self: High School Graduate			0.006***		0.006***
			(0.001)		(0.001)
Self: At least Some College			0.028***		0.028***
			(0.004)		(0.004)
Observations	82810258	82810258	82810258	82810258	82810258
Mean of Dep. Var.	0.011	0.011	0.011	0.011	0.011

Notes: Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, a quadratic in age, and father's age. Columns (2) through (5) include indicators for father being between the 50^{th} and 75^{th} percentile of income, and between the 75^{th} and 90^{th} percentile of income as independent variables. The omitted categories are below median income and less than high school eduction. Source: 1940 Historical Census Data, USPTO patent records.

Column 3 includes the child's own education. The effect of parental income disappears, instead loading on the child's own education. Those with at least some college education were 0.028 percentage points more likely to become an inventor than those with less than a high school degree. These patterns are repeated in columns 4 and 5. Column 4 shows the effect of father's education—as a proxy for income when the child was young—is highly statistically significant but this effect largely disappears when including the child's own education in column 5. One explanation of these results is that parental income affected the probability of becoming an inventor largely through its impact on children's access to education.[13]

[12] In all regressions with father's income, we include indicators for the father being between the the 50^{th} and 75^{th} percentile of income, and between the 75^{th} and 90^{th} percentile of income as independent variables, but do not report their coefficients due to space constraints.

[13] The patterns presented here can not be explained by differences in occupation choice, as shown by Appendix

While Table 2 focuses on the extensive margin—the characteristics of those becoming inventors—Table 3 considers the relationship between an inventor's background and his productivity on the intensive margin, measured by the number of career patents he generates. Column 1 shows a weak and statistically insignificant positive effect of the father being an inventor. Contrary to the results of Table 2, columns 2 and 4 do not exhibit a strong effect of father's income or father's education. Columns 3 and 5 introduce the child's own education. Again, college attendance is strongly correlated with long run inventiveness, with college-educated inventors receiving 0.286 more patent grants than their counterparts with less than a high school education on average, an increase of 21% of a standard deviation. These patterns are robust to measuring inventiveness by an inventor's log career citation counts (see Appendix Table A-9). The most highly-educated inventors tended to be the most productive.

Table 3: Individual Background and Career Patent Counts

	(1)	(2)	(3)	(4)	(5)
Father Inventor	0.219	0.236	0.098	0.213	0.061
	(0.703)	(0.662)	(0.655)	(0.673)	(0.673)
Father Income $90^{th} - 95^{th}$ %ile		-0.331	-0.320		
		(0.232)	(0.233)		
Father Income 95^{th} %ile and above		0.026	-0.041		
		(0.199)	(0.193)		
Father: High School Graduate				0.060	-0.044
				(0.113)	(0.117)
Father: At least Some College				0.188	0.064
				(0.142)	(0.128)
Self: High School Graduate			0.039		0.038
			(0.039)		(0.039)
Self: At least Some College			0.286***		0.286***
			(0.044)		(0.044)
Observations	9032	9032	9032	9032	9032
Mean of Dep. Var.	1.581	1.581	1.581	1.581	1.581
S.D. of Dep. Var.	1.365	1.365	1.365	1.365	1.365

Notes: Table reports coefficients from a regression in which the dependent variable is log career patent counts for the sample of inventors in our matched sample. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, a quadratic in age, and father's age. Columns (2) and (3) include indicators for father being between the 50^{th} and 75^{th} percentile of income, and between the 75^{th} and 90^{th} percentile of income as independent variables. The omitted income category is below median income, and we omit an indicator for the individual having less than a high school education. *,** ,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level. Source: 1940 Historical Census, USPTO patents.

Two findings emerge when taking Table 2 and Table 3 together. First, the importance of education holds both at the extensive and intensive margins, which is consistent with a human capital explanation of invention. Second, both father inventor status and parental income matter on the extensive margin but not on the intensive margin, which suggests that the existence of credit constraints might have undermined inventiveness. This second finding is related to a long line of research in the family firm and management practice literatures, showing that privileged

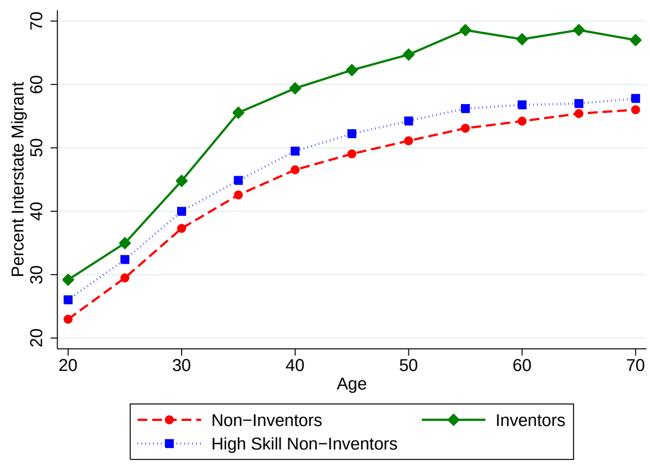
Table A-8, which introduces occupation fixed effects to every column of Table 2.

access to career paths (e.g., inherited CEO roles) is associated with under performance (e.g., Perez-Gonzalez (2006), Bloom and Van Reenen (2007), Caselli and Gennaioli (2013)).

Fact 3 ***Inventors were more likely to have migrated from their state of birth. They moved to states that were more conducive to innovation.***

Individuals migrate in order to seek better job prospects in their destination state. This argument may apply particularly strongly for inventors, since environmental factors shift both the costs and benefits of innovation. The example of Thomas Edison illustrates this point. Not only did he stand to gain more from marketing his inventions in the larger market of New Jersey and New York but he also benefitted from the larger supply of skilled labor and financial development there. When inventors systematically move to such places, this generates spatial concentration giving rise to agglomeration externalities.

Figure 9: Interstate Migration Rates by Age



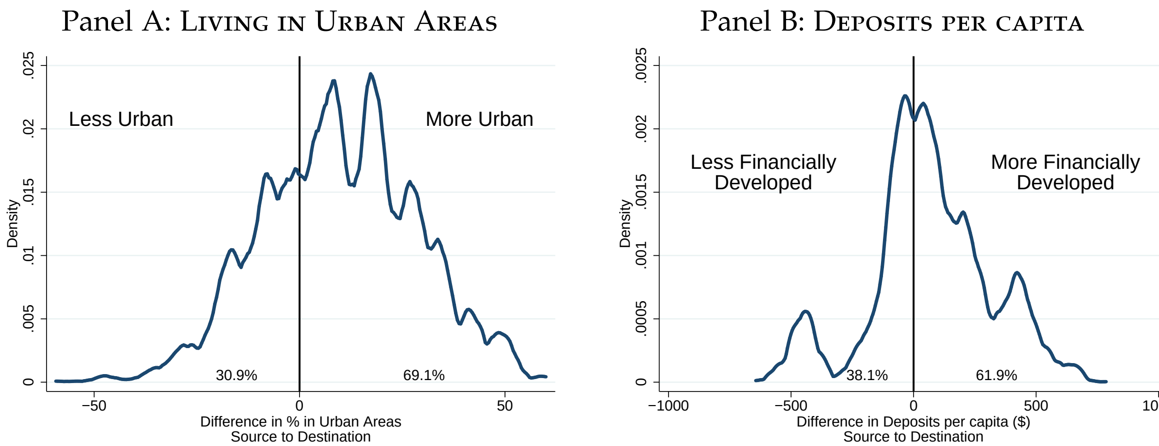
Notes: Figure plots interstate migration rates by age of individual for the population of high skill individuals. An individual is defined to be an interstate migrant if their birth state is different from their current state of residence. Each point represents a 5-year forward-looking bin. For example, the point at age 20 measures the average migration rate for 20 to 25 year-olds. Figure uses data averaged across the four census years for which we have occupation information: 1880, 1920, 1930, and 1940. Source: 1880, 1920-1940 Historical Census Data, USPTO patent records.

Figure 9 confirms that Edison's example is representative of the broader inventor population. The figure shows that inventors were most likely to move after the age of 35: the beginning of their most innovative period according to Figure 7. The high migration rate for inventors does not simply reflect their higher average skill level. Highly skilled individuals in non-inventor occupations migrated significantly less than do inventors.[14]

Conditional on moving to a new location, where did inventors go? To answer this question, Figure 10 plots the characteristics of geographic origin and destination amongst inventors who move across state lines in our matched dataset of inventors to the Census. We highlight two key variables strongly correlated with innovation in Figure 5. Panel A shows that inventors generally

[14] The difference between inventors and high skill non-inventors is statistically significant at the 10% level for 20-25 year olds, at the 5% level for 25-30 year olds, and at the 1% level thereafter.

Figure 10: To Where did Inventors Move?



Notes: Figure shows distribution of difference in characteristic between source and destination states for migrant inventors. The leftmost percentage on each graph corresponds to the share of migrant inventors who move to locations with a lower value of the variable of interest than their source state, while the rightmost percentage corresponds to the share that move to locations with a higher value of this variable. For instance, 30.9% of inventors move from a more urban state to a less urban state, leaving 69.1% of inventors to move to more urban states. Source: 1860, 1940 Historical Census Data, FDIC, USPTO patent records.

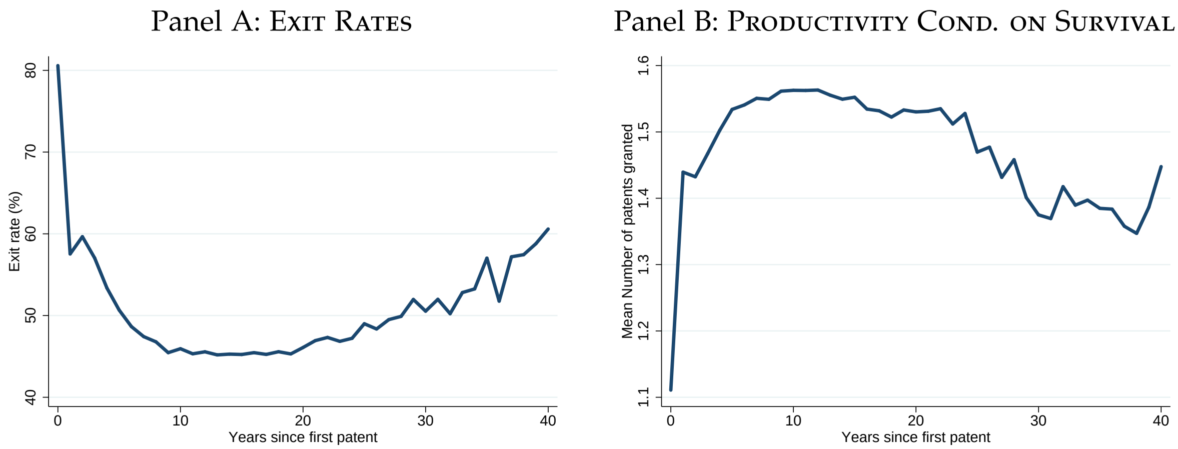
moved from less to more urbanized regions. Panel B shows that inventors moved toward regions where deposit ratios were higher, suggesting that access to finance could have played a role in their migration decisions. Both of these figures suggest that inventors generally migrated to regions whose characteristics were well-suited to innovation.

Fact 4 ***Inventors were positively selected through exit early in their careers, increasing the average productivity (conditional of survival) of a cohort of inventors. However, they were less productive and more likely to exit late in their careers.***

If inventors approximate the life cycle of firms, some should enter, develop and succeed whereas others should fail and exit—entrepreneurial churn is an essential feature of a well-functioning innovation sector (Haltiwanger, 2012). Figure 11 plots the career cycle of inventors using the universe of inventor data, as opposed to just the inventor data matched to the Census. Panel A plots the exit rate for inventors over their life cycle, where an inventor is said to have exited in period t if they file no successful patent applications in every period $t' > t$. Panel B plots the average number of patents conditional on survival for inventors over their tenure in the data. In both panels, the horizontal axis plots the number of years since the inventor filed his first successful patent application.

The figure reveals both similarities and differences with the life cycle dynamics of firms. We find evidence for both positive selection through exit, and eventual obsolescence of inventors. Panel A of Figure 11 shows that inventor exit rates exhibit a U-shape, while Panel B shows that the number of patents conditional on survival has an inverted-U shape over the life cycle. Positive selection occurs early in an inventor's career, where low productivity inventors stop applying for patents. This yields a decreasing exit rate and increasing average productivity over the average inventor's life cycle. In later years of life, however, skill obsolescence and old age set in, reducing

Figure 11: Inventors' Career Dynamics



Notes: In each panel, the horizontal axis plots the number of years since the inventor's first patent application. Panel A plots the exit rate for inventors over their life cycle, where an inventor is said to have exited in period t if they file no patent applications in every period $t' > t$. Panel B plots the average number of patents conditional on survival for inventors over their tenure in the data. Source: USPTO disambiguated inventor data 1920-2006, constructed by the authors using the algorithm of Li et al. (2014).

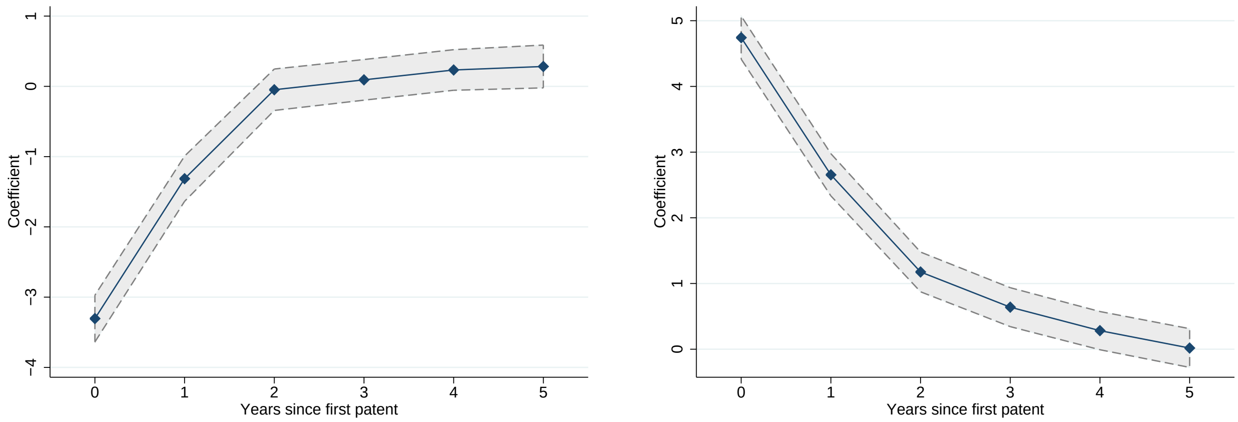
inventor productivity, and increasing exit rates. In the limit, biological constraints ensure that the inventor exit rate converges to one. Since new inventors can displace the ideas of older inventors, these life cycle results are consistent with creative destruction.

Fact 5 ***The patents of new inventors received more citations on average, and were more likely to be in the top decile of the citation distribution.***

Next, we investigate the quality of the patents granted to inventors over their career cycle. We proxy a patent's quality and influence using citation counts, adjusted following the method of (Hall et al., 2001). Figure 12 plots various moments of the patent quality distribution measured each year of an inventor's career, conditional on survival. Panel A plots the probability that a patent applied for t years after the inventor's first successful patent application lies in the bottom quartile; Panel B repeats the same exercise with the top quartile of citations received. Specifically, the figure plots coefficients from a patent-level regression in which the dependent variable is an indicator for whether the patent lies in a particular citation quartile, and the independent variables are indicators for whether a patent application came t years after the inventor's first patent, as well as individual and technology-year fixed effects. Both panels show that patents granted to new inventors are more likely to be highly cited than patents granted to inventors with a long record of patenting, mirroring the dynamics of innovative firms found in the previous literature. Patent applications in the first year of an inventor's inventive tenure are 4.74 percentage points more likely to lie in the top quartile of patent citations, and 3.3 percentage points less likely to be in the bottom quartile than are patents granted 6 or more years after the inventor's first patent, conditional on individual and technology-year fixed effects. These plots are especially striking since they are conditional on survival, given that Figure 11 shows positive selection among inventors who continue to innovate over a long career.

Figure 12: Patent Quality over an Inventor's Life Cycle

Panel A: Pr{1st Quartile Patent} — Panel B: Pr{Fourth Quartile Patent}



Notes: Figure plots regression coefficients from an OLS regression of the panel title on indicators for whether a patent application came *t* years after the inventor's first appearance in the patent data. All regressions include individual and technology-year fixed effects. Grey bands indicate 95% confidence interval around point estimates, using standard errors which are clustered at the technology class-year level. Source: USPTO disambiguated inventor data 1920-2006, constructed by the authors using the algorithm of Li et al. (2014).

Table 4 repeats the same analysis in a regression framework. Column 1 regresses the log number of citations on an indicator that is equal to 1 if the patent is applied for within the first two years of an inventor's career and 0 otherwise. It shows that patents obtained early in the career are of higher quality, receiving 7.7% more citations on average. Columns 2 to 5 replace the dependent variable with an indicator equal to 100 if the patent belongs to the relevant citation quartile, and 0 otherwise. Again, on average, we see that inventors produce more influential work early in their career.

Table 4: Panel Relationship between Entry and Patent Quality

	Log Citations	Patent in quartile (coefficients sum to 0):			
		First	Second	Third	Fourth
	(1)	(2)	(3)	(4)	(5)
Patent granted in first two years of inventor career	0.077***	-1.791***	-1.087***	0.350***	2.528***
	(0.002)	(0.093)	(0.092)	(0.096)	(0.091)
Inventor Fixed Effects	Y	Y	Y	Y	Y
Class $\times$ Year Effects	Y	Y	Y	Y	Y
Observations	4290376	4765684	4765684	4765684	4765684

Notes: Table reports regression coefficients from an OLS regression of log citations and whether a patent was in a particular quartile of the citation distribution on an indicator for whether a patent was granted in the first two years of a career. All regressions include individual and technology-year fixed effects. Standard errors are clustered at the technology class-year level. Source: USPTO disambiguated inventor data 1920-2006, constructed by the authors using the algorithm of Li et al. (2014).

Fact 6 ***Successful patentees had substantially higher labor income, and produced higher quality inventions. For younger inventors, future productivity predicts current income whereas for older inventors, income is predicted by both past and future productivity.***

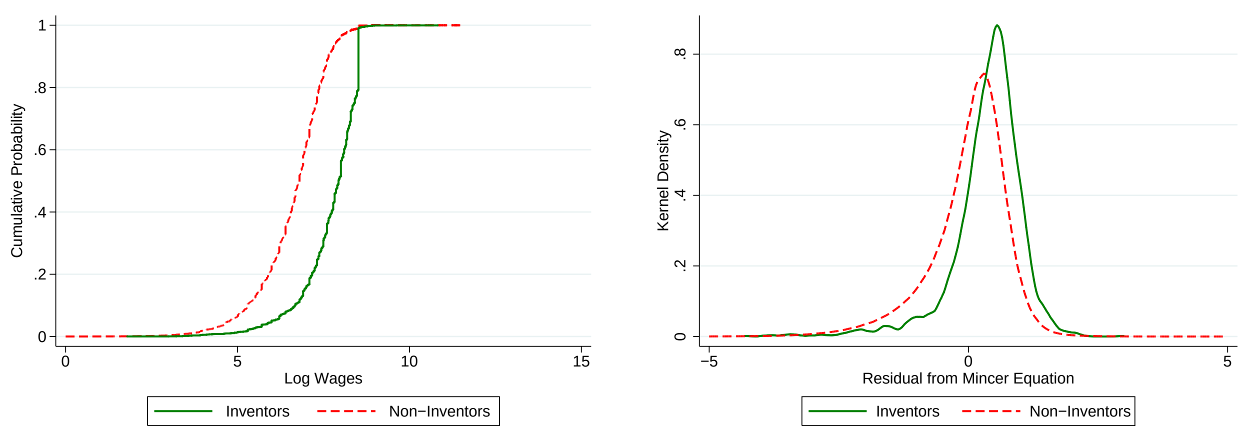
We now turn to the private returns to innovation. Schmookler (1966) argued that the expectation of pecuniary gain was implicit to most inventors' careers, citing Thomas Edison whose motivations were largely commercial. Because our wage data come from the 1940 Census two caveats are important to consider. First, labor income is not recorded for all observations in the Census;[15] and second, labor income itself provides only a partial measure of the total financial returns to innovation. The discovery of new inventions may permit individuals to start their own business and earn a return on new capital assets. Non-wage factors, which are unmeasurable to us, may be an important benefit of self-employment (Hurst and Pugsley (2011)).

Based on the wage data we do observe, Figure 13 plots the distribution of wage income for inventors and non-inventors. Panel A plots the unconditional CDF of log wage income for both groups. Unsurprisingly, inventors have relatively high incomes. Indeed, the inventors' income distribution first order stochastically dominates that of non-inventors.

Figure 13: The Distribution of Labor Income by Inventor Status (1940)

Panel A: Unconditional Distribution

Panel B: Conditional on Observables



Notes: Figure plots the distribution of the natural log of wage income for inventors and non-inventors, as reported in the 1940 census. Many individuals report 0 wages, and are excluded from this plot. Solid green lines plot the distribution of inventors' wages, while dashed red lines plot the distribution of non-inventors' wages. Panel A plots the unconditional CDF of log wages. Panel B plots the density of log wages residualized against observable characteristics. Specifically, it plots the distribution of residuals from a regression in which the dependent variable is log wages, and includes controls for race, education, sex, international migrant status, residence state fixed effects, occupation fixed effects, and a quadratic in age. Source: 1940 Historical Census Data, USPTO patent records.

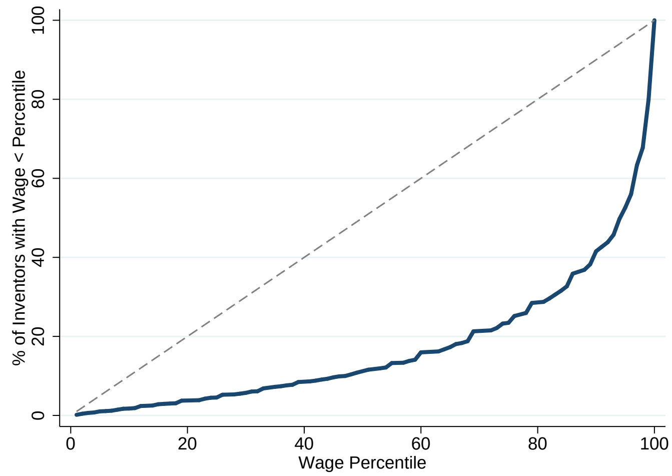
This result is expected given that inventors were better-educated, higher-skilled, and lived in more urban states than non-inventors. Panel B therefore plots the distribution of wages for inventors and non-inventors after conditioning on observables. Specifically, we regress an individual's log wages on race, education, sex, international migrant status, residence state fixed effects, occupation fixed effects, and a quadratic in age. We then plot the distribution of residuals from this regression for inventors and non-inventors. Even after controlling for all observable characteristics, inventors have higher wage incomes throughout the distribution.

Figure 14 further illustrates the inventor wage premium by showing the share of inventors with wage income *below* each wage percentile. The 45-degree line represents perfect equality in the income distribution where inventor wages mirror wages in the general population. Instead we see that the inventor wage observations are distributed well-below the diagonal. For example,

[15] We drop all those individuals who report a wage income of zero from this analysis.

only 10.8% of inventors have wages below the 50th percentile of the income distribution and 23.4% have wages below the 75th percentile. Meanwhile, fully 58.5% of inventors have incomes in the top decile of earnings.

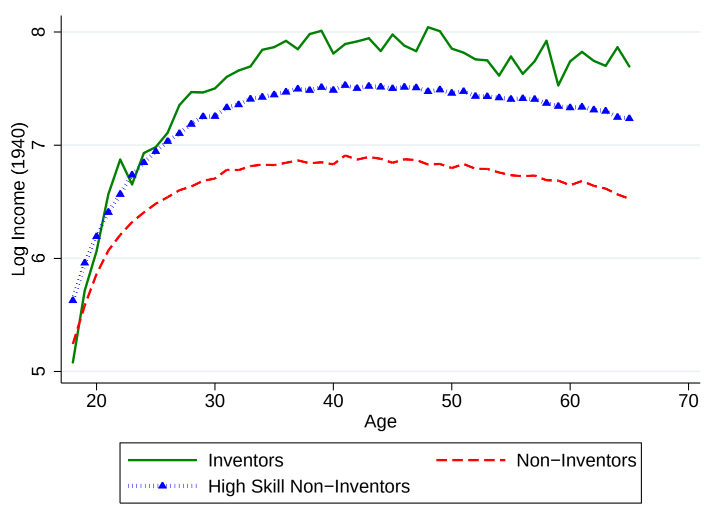
Figure 14: Share of Inventors with Incomes Below Each Income Percentile



Notes: Figure plots the distribution of wages by income percentile for inventors relative to the general population. The 45-degree line represents perfect equality in the income distribution. Source: 1940 Historical Census Data, USPTO.

Inventors also have a steeper life cycle profile of wages. Figure 15 plots the average life cycle of log earnings for inventors, non-inventors, and non-inventors in high skill occupations. This figure is constructed from the cross-section of individuals at each age. Inventors have higher earnings throughout their life cycle than non-inventors and high-skilled individuals. Table A-10 in the Appendix shows that the difference between the wages of inventors and high-skill non-inventors is statistically significant at the 1% level from the age of 19 onwards. These figures support the idea that invention was a key labor income differentiator.

Figure 15: The Life Cycle of Earnings by Inventor Status

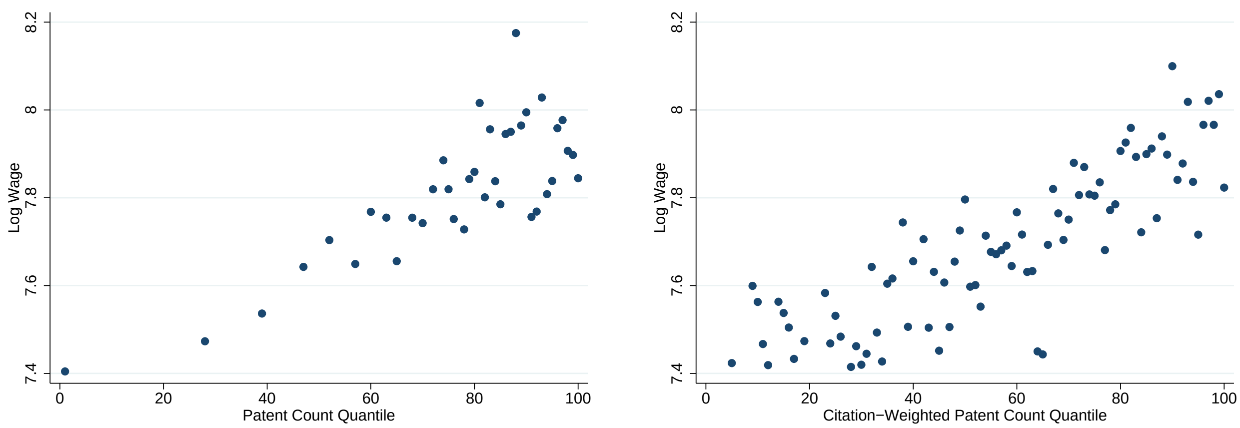


Notes: Figure plots the evolution of log average wage income over the life cycle. The solid green line plots the evolution of inventors' wage income, while the dashed red line plots the wage evolution of the universe of non-inventors. The dotted blue line plots the life cycle of all high skill non-inventors. Source: 1940 Historical Census, USPTO patents.

If the returns to invention reflected pecuniary gains from technological development we would also expect to observe a correlation between labor income and the quality of patents, as predicted by our model. The data shows a strong correlation between the quality of an inventor's patent portfolio and log wages. Panel A of Figure 16 plots the relationship between the number of patents an inventor files over his lifetime and log average wages. Panel B mirrors Panel A, except the horizontal axis now weights each patent in an inventor's portfolio by the number of citations the patent receives. Both panels exhibit a robust positive relationship between inventor productivity and log wages, suggesting that the higher-quality inventors were compensated for their inventions.

Figure 16: The Relationship between Innovative Productivity and Wages

Panel A: Unweighted Patent Counts | Panel B: Citation-Weighted Patent Counts



Notes: Figure plots the relationship between log average wages and the quantile of inventive activity, conditional on being granted at least one patent. The median inventor is granted only 3 patents. Thus the unweighted patent count has relatively few percentile points at the low end of the distribution: the first 24% of the distribution is contained in the first percentile data point. Source: 1940 Historical Census Data, USPTO patent records.

There are two possible explanations for this positive relationship between wages and inventor productivity. First, an inventor may simply be more productive as a result of his past inventions. Alternatively, if invention is a signal of underlying worker type, an employer may pay an inventor more of a financial premium in anticipation of future productivity.

To disentangle these two effects, we regress log wages, measured in 1940, on an inventor's innovative activity both before and after 1940. If the current productivity effect dominates, we would expect pre-1940 innovation to have a strong effect on wages. However, if the anticipation effect dominates, forward-looking innovative activity should predict an inventor's wages, so long as employers correctly anticipate an employee's future productivity. The results are reported in Table 5. Each regression controls for inventor demographics, education, and state. To account for differences between young and old inventors, the table reports coefficients estimated for the set of inventors both above and below the age of 35.

One might expect the anticipation effect to be stronger for young inventors who have a longer career ahead of them at the point in time that they enter the most productive part of their careers (see Figure 7). This prediction is borne out in the data. Young inventors see large gains from future innovations, regardless of whether innovation is measured using citation-weighted or un-

Table 5: What determined inventor income? Regressions of log wages on innovation measures

	Age: Under 35		Age: Over 35	
	(1)	(2)	(3)	(4)
Log Patents Pre-1940	-0.022		0.060***	
	(0.018)		(0.014)	
Log Patents Post-1940	0.087***		0.040***	
	(0.016)		(0.011)	
Log Citations Pre-1940		0.002		0.030***
		(0.009)		(0.007)
Log Citations Post-1940		0.039***		0.030***
		(0.010)		(0.008)
Observations	1602	1602	4458	4458
R-squared	0.482	0.480	0.302	0.302
Mean of Dep. Var.	7.275	7.275	7.765	7.765
S.D. of Dep. Var.	0.927	0.927	0.781	0.781

Notes: Table presents estimated coefficients from a regression of log wages on innovation measures. We restrict our attention to the sample of inventors matched to the 1940 census. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, occupation skill level, education and a quadratic in age. *,** ,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level. Source: 1940 Historical Census Data, USPTO patent records.

weighted patent counts. This suggests that the anticipation effect is strong for them. Meanwhile, there is no statistically distinguishable difference between the anticipation and past productivity effects for older inventors. At this stage, it is worth reiterating the caveat that our data only contain wage income information; older inventors with many patents may see unmeasured benefits from capital income or entrepreneurship.

5 Economic Growth, Inequality and Social Mobility

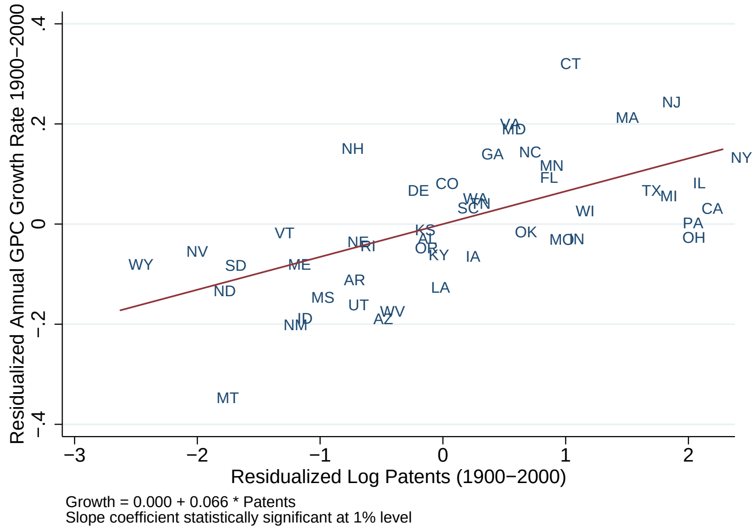
As suggested by our theoretical framework in Section 3 the micro-level results we have presented so far should also be linked to macro-level outcomes. Recall that in the model, technological innovations make production workers more productive. The long-standing endogenous growth literature builds on the premise that long-run growth is driven by innovation and technological progress. But even though a large literature has studied the empirical determinants of macro-level economic growth (e.g., Barro (1991)), to our knowledge no study has documented a causal empirical relationship between innovation and growth for the U.S. over the long run.

We use state-level patents as a proxy measure of innovation. While not all inventions are patented and unpatented process inventions like the assembly line can exert a powerful influence on economic growth, patents do provide a broad indicator of innovative activity. Furthermore, although the technology embodied in an individual patent might diffuse with some time lag, the spillovers created during its production should be highly localized. Jaffe et al. (1993) and Thompson (2006) both find evidence of localized knowledge spillovers from U.S. patents. Localization should positively impact regional economic growth, as inventors interact in close geographic proximity and learn from one another when developing new inventions.

Fact 7 ***More inventive states and sectors grew faster on average.***

Figure 17 shows the basic correlation between patents and economic growth is strongly positive. To account for the initial heterogeneity in income levels, we plot variables residualized against 1900 log GDP per capita.

Figure 17: Innovation and Long-run Growth: U.S. States between 1900-2000



Notes: Figure plots the total number of patents granted to inventors in each state between 1900 and 2000 on the horizontal axis, and the annualized growth rate in state GDP per capita between 1900 and 2000 on the vertical axis. Both horizontal and vertical axes plot the variables of interest residualized against 1900 log GDP per capita, to control for conditional convergence. Source: BEA Historical Regional Economic Accounts, and Klein (2013).

Table 6 reports coefficients from growth regressions controlling for the long-run effects of initial conditions and population density. The dependent variable in these regressions is the annualized growth rate in state-level GDP per capita between 1900 and 2000.

Table 6: Innovation and Long Run Growth: U.S. States between 1900-2000

	Annualized Growth Rate		DHS Growth Rate	
	(1)	(2)	(3)	(4)
Log Patents	0.066***	0.054***	0.031***	0.026***
	(0.013)	(0.012)	(0.008)	(0.007)
Initial GDP per Capita	-0.877***	-0.891***	-0.324***	-0.330***
	(0.036)	(0.036)	(0.025)	(0.026)
Population Density		1.145*		0.517*
		(0.588)		(0.304)
Observations	48	48	48	48
Mean Growth	2.154	2.154	1.552	1.552
Std. Dev. of Growth	0.417	0.417	0.159	0.159

Notes: Table reports estimated coefficients from a regression in which the dependent variable is the state-level annualized growth rate in real GDP per capita from 1900-2000. White heteroskedasticity robust standard errors reported in parentheses. DHS growth rate refers to the growth rate measure as proposed by Davis, Haltiwanger, and Schuh. Output data provided by Klein (2013) and the Bureau of Economic Analysis. *,**,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level.

We find that the log of patents granted between 1900 and 2000 had a consistently positive and statistically significant effect in columns 1 and 2. These results are robust in columns 3 and 4 to measuring the growth rate using the approach established by Davis et al. (1996) in the employment literature that corrects for any potential bias associated with transitory shocks to growth and mean reversion.[16]

The economic magnitude of these estimates is especially informative. Consider two states: a low innovative state Mississippi (at the 10th percentile) and high innovative state New Jersey (at the 90th percentile). Assume NJ and MS had the same initial GDP per capita in 1900 and identical population densities. Our estimated coefficients imply that the gap between NJ and MS would have increased dramatically. By the end of the century, NJ would be 26% richer than MS just because of the differences in their innovativeness.

Instrumental Variables

We now attempt to identify a causal effect in our growth regressions using contracts for wartime technological development as an instrument for innovation disbursed by the Office of Scientific Research and Development (OSRD). A brief survey of the institutional setting along with quantitative tests lends support to the credibility of our instrumental variables approach.

The OSRD was established under an Executive Order from President Roosevelt in June 1941, and operated until its termination in December 1947. Headed by Vannevar Bush at the Carnegie Institution of Washington, the OSRD was responsible for major innovations that had an impact in wartime and beyond, including miniature electronics like the proximity fuse, navigation systems, solid fuel rockets, detonators and most famously the basic science used in the Manhattan Project. Because of its significant impact, the OSRD spurred federal involvement in the development of U.S. science and technology in the postwar years (Stephan, 2014).

The OSRD did not operate laboratories of its own; rather it contracted out the development of inventions. This reflected a new way of mobilizing public funding for the development of scientific resources. During World War I scientists had worked at rudimentary laboratories established by the government on an *ad hoc* basis, and there was a long-standing concern among scientists that federal involvement in their activities would threaten creativity and intellectual independence. As Mowery (2010, p.1227) comments, "the contractual arrangements developed by the OSRD during World War II allowed the office to tap the expanded range of private sector and university scientific and engineering capabilities that had developed during the interwar period."

However, the OSRD did not know *ex ante* which firms or academic institutions would be successful because "the OSRD had long insisted that it was not working on materials or methods of wide use in industry" (NAS, 1964, p.28). In fact, due to this uncertainty, the OSRD sometimes contracted with multiple entities to solve the same problem. The OSRD spent $450 million in total, about six and a half times the federal budget for science in 1940. Around this time universities had been spending about $50 million on research of which around $6 million was funded by the federal government to support mostly agriculture-related research (Payne, 1992,

[16]Figure A-6 in appendix D shows that this strong positive relationship between long run growth and innovation holds for historical output calculated using the methodology of Martin (1939).

p.145). The OSRD created a large boost to firm-level R&D. For example, Radio Corporation of America invested heavily at its plants in Indiana and New Jersey (Chandler, 2001 p.27-28).

We collected data on all contracts granted by the OSRD. We observe 1,717 contracts across 39 U.S. States. The coverage of the OSRD contracts is wide. For example, Iowa State College received 10 contracts and the University of New Mexico received 7 contracts. Firms and academic institutions in the state of New York accounted for 30 percent of the total with the next largest concentrations of contracts being in Massachusetts (13 percent) and Pennsylvania (11 percent). The mean number of contracts per firm/academic institution was 4.3 and the median was 1. The most awarded private firm was the Western Electric Company with 107 contracts. The most prolific university was MIT which was granted 89 contracts.

Our strategy requires that these contracts were correlated with innovation, uncorrelated with omitted determinants, and only influenced state growth rates through their effect on innovation. Note that if the OSRD contracted with only the best firms or academic institutions (which it did not), this would not be a violation of the exclusion restriction, so long as initial location decisions were orthogonal to a state's future growth rate.

Table 7 reports coefficients from a regression of post-war state-level growth in GDP per capita for a four decade time horizon (1947-1987) on state innovation levels immediately following World War II (1945-1950).[17] In columns 1 and 2 we report OLS estimates controlling for the long-run effects of initial GDP per capita and population density respectively. The corresponding IV estimates in columns 3 and 4 include these controls and a control for long run past growth to address the potential confound that contracts were just awarded to high growth areas. Column 5 reports the first stage regression of log patents on the number of OSRD contracts.

The first stage relationship is strongly positive and interesting in its own right. Barro (1981) and Field (2008) show that general wartime spending had little impact on economic growth and may have even crowded out private sector investment. Fishback and Cullen (2013) find that "growth in per capita measures of economic activity [to 1958] showed little relationship with per capita war spending" and Jaworski (2015) finds little effect of wartime spending on subsequent growth rates in the U.S. South.

These studies suggest that our use of OSRD contracts as an instrument will not be invalidated by any correlated contemporaneous response of GDP per capita to other forms of government-spending. OSRD contracts were targeted towards innovation, which we would expect to be related to long run growth, whereas more general government contract spending on combat-related equipment, like aeroplanes and tanks, or incidentals, such as clothing, was not.

The OLS coefficients from Table 7 are broadly similar to the IV estimates, if also somewhat smaller. In column 3 the IV estimate is about 26% larger than the corresponding OLS estimate in column 2, which would be consistent with the OSRD financing innovations that had an especially sizeable impact on economic growth, under a local average treatment effect interpretation. U.S. technological leadership was tightly linked to economic growth in the post World War II years (Nelson and Wright, 1992). Equally, because we do not take into account the positive cross-state

[17] The results are similar if we instead study the effect of patenting during the war, between 1940 and 1945. Additionally, Table A-7 in Appendix D shows that the patterns are robust to measuring growth rates using the methodology of Davis et al. (1996).

Table 7: Innovation and Long Run Growth: U.S. States between 1947-1987

	Annualized Growth Rate				1^{st} Stage
	OLS	OLS	IV	IV	OLS
	(1)	(2)	(3)	(4)	(5)
Log Patents (1945-1950)	0.123***	0.101***	0.127***	0.082**	
	(0.028)	(0.031)	(0.038)	(0.039)	
OSRD Contracts					0.698***
					(0.083)
Log GDP per Capita (1945)	-1.655***	-1.688***	-1.738***	-1.511***	0.250
	(0.148)	(0.148)	(0.147)	(0.125)	(0.638)
Population Density (1945)		1.064	0.798	0.820	0.574
		(0.652)	(0.575)	(0.588)	(2.291)
1900-1940 GDP/cap. Annual Growth Rate				0.146**	0.391*
				(0.067)	(0.214)
Observations	48	48	48	48	48
Mean Growth	2.501	2.501	2.501	2.501	6.698
Std. Dev. of Growth	0.439	0.439	0.439	0.439	1.502
F-Statistic					66.126

Notes: Table reports estimated coefficients from a regression in which the dependent variable is the state-level annualized growth rate in GDP per capita from 1947-1987. White heteroskedasticity robust standard errors reported in parentheses. The IV estimates are two-stage least squares estimates using the number of OSRD contracts in each state during World War II. *,** ,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level, respectively.

spillovers from successful innovations, our state-level estimates (and IV estimates) will likely understate the aggregate relationship between innovation and economic growth.

Table 8: Testing for Selection Effects in OSRD Contracts

	t	
	1935-40	1930-1935
	(1)	(2)
Real GDP Growth (t)	0.017	0.230
	(0.100)	(0.139)
Real GDP Growth ($t-1$)	0.118	-0.302*
	(0.150)	(0.151)
GDP per Capita (t)	2.179*	1.240*
	(1.140)	(0.658)
Population Density	11.546**	12.235**
	(5.120)	(4.565)
Observations	48	48

Notes: Table reports coefficients from a regression in which the dependent variable is the number of OSRD contracts in each state during World War II and the independent variables are pre-trend growth rates, population density, and beginning of period GDP per capita. We consider growth rates from 1935-1940 (t) and 1930-1935 ($t-1$) in column 1, while in columns 2 we consider growth rates from 1930-1935 (t) and 1925-1930 ($t-1$). White heteroskedasticity robust standard errors reported in parentheses. *,** ,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level. The results do not change if we consider Davis, Haltiwanger, and Schuh growth rates. Source: Bureau of Economic Analysis, USPTO patent records.

Finally, to evaluate the validity of the exclusion restriction, we provide quantitative tests of the instrument in Table 8. Specifically, we check if contract allocation is correlated with pre-trend growth. While we control for this source of endogeneity explicitly in column 4, Table 7 using a variable for annualized state-level GDP per capita growth between 1900 and 1940 in Table 8 we focus on more recent periods between 1930 and 1940. We do not find a statistically significant effect of pre-period growth rates on contracts for these adjacent years.

Fact 8 ***Broad measures of income inequality (90/10, Gini) were negatively correlated with innovation, however, top-1 income share had a U-shaped relationship with innovation.***

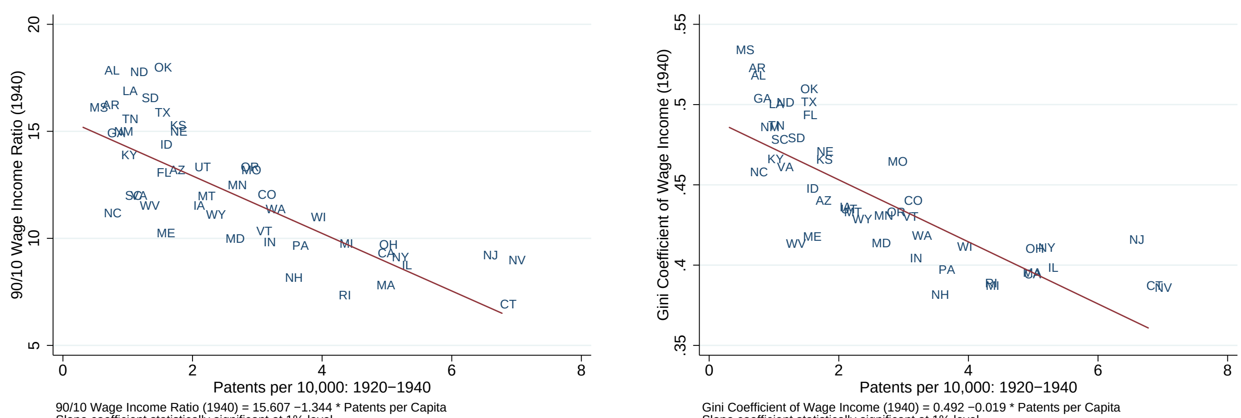
Our theoretical framework in Section 3 also suggests that innovation generates turn-over in society. The impact of innovation on inequality and social mobility is especially relevant because the existing empirical literature is divided on the topic. Aghion et al. (2015a) examine modern U.S. data finding a positive causal effect of innovation-led growth on top incomes shares at the state-level. However, they also find some sensitivity to measurement. The relationship between inequality and patenting becomes much weaker at different thresholds like the top 10% share, and they find a negative relationship when using the Gini coefficient, which considers all parts of the income distribution rather than just the top share. By contrast Jones and Kim (2014) shows theoretically that if innovations come from new entrants, the relationship between inequality and innovation could be negative.

Our results in Figure 18 generally point to a negative relationship between income inequality and inventiveness. The vertical axis plots the state-level 90/10 ratio and Gini coefficient as measured in the 1940 Census, while the horizontal axis plots backward-looking average patents per capita between 1920 and 1940. Both of these measures of inequality are strongly negatively associated with regional inventiveness.

Figure 18: Relationship between Wage Income Inequality and Inventiveness

Panel A: Ratio of 90^{th} to 10^{th} Percentile

Panel B: Gini Coefficient



Notes: Figure plots the relationship between average patents per 10,000 residents between 1920 and 1940, and the state-level wage income inequality observed in the 1940 census. Panel A measures income inequality with the ratio of the 90^{th} percentile to the 10^{th} percentile of income, while panel B uses the Gini coefficient as its measure. Delaware excluded as an outlier for visibility. Source: 1940 Historical Census Data, USPTO patent records.

Table 9 reports the results from a state-level regression of 1940 wage income inequality on average patents per capita between 1920 and 1940, and the state's occupation mix. All independent variables in the regression are standardized to have zero mean and unit standard deviation. Column 2 shows that increasing the number of patents per capita by one standard deviation is associated with a decline in the 90/10 ratio of 0.28 (= 0.828/2.98) standard deviations, conditional on the state's occupation mix.

Table 9: Wage Income Inequality and Innovation

Dependent Variable:	90/10 Ratio		Gini Coefficient	
	(1)	(2)	(3)	(4)
Av. Patents per Capita 1920-1940	-2.210***	-0.828**	-0.030***	-0.010
	(0.358)	(0.343)	(0.006)	(0.007)
% Agricultural Occupation (1940)		1.777***		0.020***
		(0.343)		(0.006)
% Manufacturing Occupation (1940)		-0.086		-0.012***
		(0.216)		(0.003)
Observations	48	48	48	48
R-squared	0.5545	0.7150	0.5239	0.7509
Mean of Dep. Var.	12.30	12.30	0.44	0.44
Std. Dev. of Dep. Var.	2.98	2.98	0.04	0.04

Notes: Table reports estimated coefficients from a regression of 1940 income inequality, measured by the ratio of the 90^{th} to the 10^{th} percentile of wage income (columns 1 and 2) and the Gini coefficient (columns 3 and 4), on the average patents per 10,000 residents between 1920 and 1940. Independent variables standardized to have zero mean and unit standard deviation. White heteroskedasticity robust standard errors reported in parentheses. *,**,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level. Source: 1940 Historical Census Data, USPTO patent records.

However, the estimated relationship between innovation and income inequality is sensitive to measurement. The top 1% income share shown in Figure 19 exhibits a non-linear, U-shaped relationship with patenting. We show this using income data in the 1940 Census and full income data constructed from Internal Revenue Service individual tax filing data from Frank (2009). The latter series will correct for any bias in the the top 1% share estimates due to the top-coding of Census income data at the $5,000 threshold.

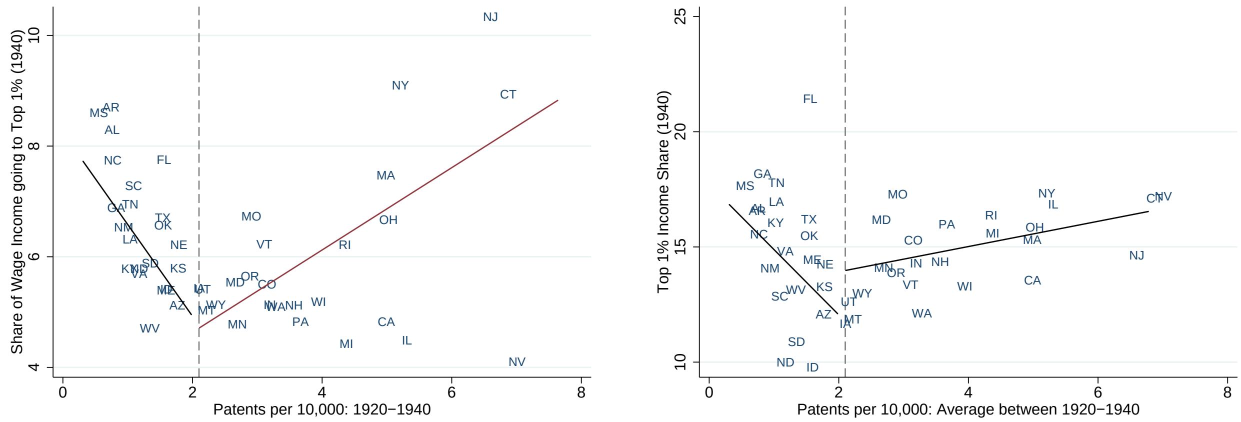
Both figures reveal a robust pattern. In the least innovative states we find a negative relationship. However, in the most innovative states such as New York, New Jersey and Massachusetts we find that more patenting was associated with more income held by the top 1%. One potentially confounding effect is the different mixes of occupations in these innovative states. For example, Philippon and Reshef (2012) show that between 1909 and 1933 skill-based wage-compensation in finance was high. Addressing this concern, Appendix Figure A-9 shows that these patterns are robust to excluding individuals who work in the financial sector.

Furthermore, Appendix Figure A-10 shows that we can reconcile the negative relationship between inequality and inventiveness shown in Figure 18 with the U-shaped relationship shown in Figure 19. When we use the top 10% income share the relationship approximates the pattern in the data we see when using the Gini coefficient or the 90/10 ratio. The U-shaped relationship only begins to emerge beyond the top 5% threshold. This suggests the very upper end of the income distribution is key to understanding the link between innovation and inequality.

Although we do not claim to show a causal link between innovation and income inequality, our analysis yields a number of important insights. First, alternative measures of inequality may yield startlingly different results. The literature has not yet reached a consensus on the economics behind the differences seen in these various measures. Second, the correlations presented here

Figure 19: Relationship between Top-1% Labor-income Share and Inventiveness

Panel A: Census Wage Income Data | Panel B: Full Income Data



Notes: Figure plots the relationship between average patents per 10,000 residents between 1920 and 1940, and 1940 state-level wage income inequality, measured by the share of income controlled by the top 1% of the state's wage earners. Delaware excluded as an outlier for visibility. Source: 1940 Historical Census Data, USPTO patent records, Frank (2009).

suggest that innovation may indeed have an important effect on income inequality. Note that our data cover a period in U.S. history when income inequality was especially high (Goldin and Katz (1999), Piketty and Saez (2003)).

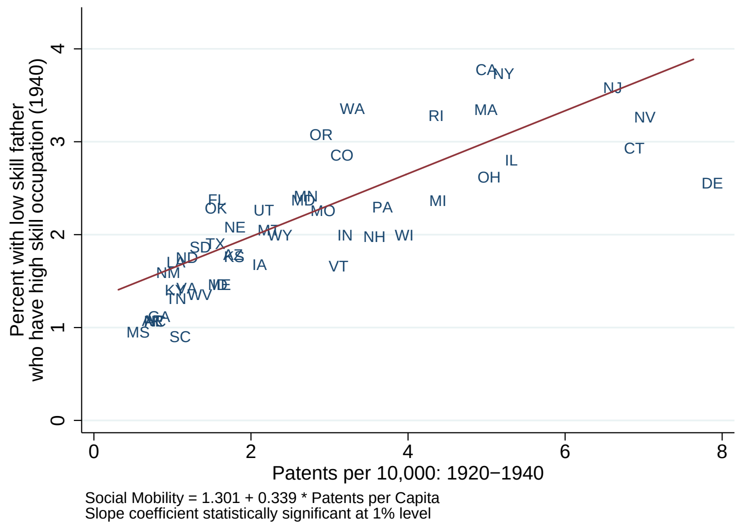
Fact 9 ***Innovation was strongly positively correlated with social mobility.***

While places with high income inequality tend to exhibit low levels of social mobility (Chetty et al., 2014), innovation might actually decrease inequality if it acts as a social elevator, as exhibited in our model. The Schumpeterian paradigm suggests that innovation allows new entrants to capture markets from old incumbents. This process of creative destruction creates churn in the economy, allowing individuals and firms with limited market shares to grow. This mechanism lies at the heart of a large class of economic models, such as Aghion and Howitt (1992), Klette and Kortum (2004), and Akcigit and Kerr (2017).

We examine the relationship between innovation and social mobility directly using our occupation data. Figure 20 shows how a state's level of social mobility in 1940 correlates with the number of patents per capita granted between 1920 and 1940. Social mobility is measured as the fraction of individuals with a low skill father, who themselves have a high skill occupation. The figure implies that more innovative regions featured more social mobility.

Of course, more innovative states may have higher social mobility for a number of reasons. For example, we have shown that innovative states tend to be more urban which may be correlated with social mobility while states with different types of economic activity (agricultural versus manufacturing) may also vary in their degree of upward mobility. To counter this concern, we regress our measure of social mobility in 1940 on average patents granted per capita between 1920 and 1940, and controls for a state's occupation mix. The percent of people in a state who work in the agricultural sector serves as a proxy for the state's degree of urbanization. Table 10 shows that the relationship between patents per capita and social mobility is positive

Figure 20: The Relationship between Inventiveness and Social Mobility



Notes: Figure plots the relationship between average patents per 10,000 residents between 1920 and 1940, and 1940 social mobility, measured by the share of those with a low-skill father who themselves have a high skill occupation. Source: 1940 Historical Census Data, USPTO patent records.

and statistically significant, even after controlling for the state occupation mix.

Table 10: % of High-skill Child given Low-skill Father

	(1)	(2)
Av. Patents per Capita 1920-1940	0.746^{***}	0.484^{***}
	(0.116)	(0.149)
% Agricultural Occupation (1940)		-0.031^{***}
		(0.011)
% Manufacturing Occupation (1940)		-0.016
		(0.019)
Observations	49	48
R-squared	0.5924	0.6844

Notes: Table reports estimated coefficients from a regression of 1940 social mobility, measured by the share of those with a low-skill father who themselves have a high skill occupation, on the average patents per 10,000 residents between 1920 and 1940. Both dependent and independent variables standardized to have zero mean and unit standard deviation. White heteroskedasticity robust standard errors reported in parentheses below coefficient. *,**,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level. Source: 1940 Historical Census Data, USPTO patent records.

These results suggest that innovation could have been a key driver of social mobility. The analysis is important in light of the finding of Long and Ferrie (2013) that America generally became less socially mobile around the turn of the twentieth century (down from its mid-nineteenth century high-point). While we do not measure changing mobility levels over time, our results do indicate that innovative places were also socially mobile places. Moreover, since Table 1 shows inventors were a small share of the total population who had a outsized effect on U.S. development, our findings underscore the need to study social movement with this important sub-group of the population in mind.

6 Conclusion

This paper presents a series of facts emerging from a major data collection exercise combining U.S. patent records with Federal Censuses between 1880 and 1940 and regional economic aggregates. The new data provide a comprehensive profile of inventions during the golden age of U.S. invention and can complement modern studies such as Aghion et al. (2015b) and Bell et al. (2015) to provide a more complete picture of inventor profiles over time and space.

Examining the drivers of innovation during this historical time period sheds light on numerous debates on innovation and long-run economic growth. For example, our evidence on the family background of inventors pinpoints an important role for education and human capital accumulation. Entry into an inventive career was increasing in father's income, but the mechanism appears to operate through improved access to education for children. Our evidence on the life cycle of invention highlights creative destruction dynamics. Inventors were positively selected early in their careers on the quality of their inventions but their productivity dropped sharply later in their careers, presumably as new entrants disrupted existing ideas.

Our micro-level evidence provides key linkages to our macro-level findings. Although we found inventors to be a small sub-group of society they had a tremendous influence on economic growth. Inventors tended to move to urban and financially developed places favorable to innovating, and they received large pecuniary payoffs from engaging in technological development. While we have shown that regional innovation yields growth, it is also related to inequality and social mobility. Establishing the background of the most effective inventors can therefore inform well-targeted innovation policies. The extent to which innovation contributes to growth, inequality, and mobility is central to determining the societal costs and benefits of technological advance.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson: 2014, 'A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration'. *Journal of Political Economy* **122**(3), 467–506.

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr: 2013, 'Innovation, Reallocation, and Growth'. National Bureau of Economic Research Working Paper #18993.

Acemoglu, D., U. Akcigit, and M. A. Celik: 2014, 'Young, Restless and Creative: Openness to Disruption and Creative Innovations'. National Bureau of Economic Research WP # 19894.

Acemoglu, D., U. Akcigit, and W. R. Kerr: 2016a, 'Innovation Network'. *Proceedings of the National Academy of Sciences* **113**(41), 11483–11488.

Acemoglu, D., J. Moscona, and J. A. Robinson: 2016b, 'State Capacity and American Technology: Evidence from the 19th Century'. Working Paper 21932, National Bureau of Economic Research.

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hémous: 2015a, 'Innovation and Top Income Inequality'. National Bureau of Economic Research Working Paper # 21247.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen: 2015b, 'Living the American Dream in Finland: The Social Mobility of Innovators'. University of Chicago mimeo.

Aghion, P. and P. Howitt: 1992, 'A Model of Growth through Creative Destruction'. *Econometrica* **60**(2).

Akcigit, U., M. A. Celik, and J. Greenwood: 2016a, 'Buy, Keep or Sell: Economic Growth and the Market for Ideas'. *Econometrica* **84**(3), 943–984.

Akcigit, U., J. Grigsby, and T. Nicholas: 2017a, 'Immigration and the Rise of American Ingenuity'. *American Economic Review, Papers and Proceedings* **107**(5), 327–331.

Akcigit, U., J. Grigsby, and T. Nicholas: 2017b, 'The Rise of American Ingenuity: Innovation and Inventors of the Golden Age'. National Bureau of Economic Research WP #23047.

Akcigit, U., J. Grigsby, T. Nicholas, and S. Stantcheva: 2017c, 'The Long-run Effects of Corporate and Personal Income Taxes on Innovation in the U.S.: Evidence from Historical Patent Data'. University of Chicago, Mimeo.

Akcigit, U. and W. R. Kerr: 2017, 'Growth Through Heterogeneous Innovations'. *Journal of Political Economy*. forthcoming.

Akcigit, U., E. Miguelez, V. Sterzi, S. Stantcheva, and S. Caicedo: 2016b, 'Dancing with the Stars: Interactions and Human Capital Accumulation'. University of Chicago Working Paper.

Alvarez, F. E., F. J. Buera, and R. E. Lucas: 2013, 'Idea flows, economic growth, and trade'. National Bureau of Economic Research WP #19667.

Banerjee, A. V. and E. Duflo: 2003, 'Inequality and Growth: What Can the Data Say?'. *Journal of Economic Growth* **8**(3), 267–299.

Banerjee, A. V. and E. Duflo: 2005, 'Growth Theory through the Lens of Development Economics'. *Handbook of Economic Growth* **1**, 473–552.

Barro, R.: 1981, 'Output Effects of Government Purchases'. *Journal of Political Economy* **89**(6), 1086–1121.

Barro, R. J.: 1991, 'Economic Growth in a Cross Section of Countries'. *Quarterly Journal of Economics* **106**(2), 407–443.

Becker, G.: 1967, *Human Capital and the Personal Distribution of Income: An Analytical Approach*. Ann Arbor: University of Michigan Press.

Becker, G. S.: 1974, 'A Theory of Marriage: Part II'. *Journal of Political Economy* **82**(2), S11–S26.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen: 2015, 'The Lifecycle of Inventors'. Harvard mimeo.

Bergstrom, T. C. and M. Bagnoli: 1993, 'Courtship as a Waiting Game'. *Journal of Political Economy* **101**(1), 185–202.

Bloom, N., M. Schankerman, and J. Van Reenen: 2013, 'Identifying Technology Spillovers and Product Market Rivalry'. *Econometrica* **81**(4), 1347–1393.

Bloom, N. and J. Van Reenen: 2007, 'Measuring and Explaining Management Practices Across Firms and Countries'. *Quarterly Journal of Economics* **122**(4), 1351–1408.

Carlino, G. and W. R. Kerr: 2015, 'Agglomeration and Innovation'. In: *Handbook of Regional and Urban Economics*, Vol. 5. Elsevier, Chapt. 6, pp. 349–404.

Caselli, F. and N. Gennaioli: 2013, 'Dynastic Management'. *Economic Inquiry* **51**(1), 971–996.

Celik, M. A.: 2015, 'Does the Cream Always Rise to the Top? The Misallocation of Talent and Innovation'. University of Pennsylvania Mimeo.

Chandler, A. D.: 2001, *Inventing the Electronic Century: The Epic Story of the Consumer Electronics and Computer Industries*. Cambridge, Mass.,: Harvard University Press.

Chetty, R., N. Hendren, P. Kline, and E. Saez: 2014, 'Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States'. *Quarterly Journal of Economics* **129**(4), 1553–1623.

Cook, L. D.: 2011, 'Inventing Social Capital: Evidence from African American inventors, 1843 – 1930'. *Explorations in Economic History* **48**(4), 507–518.

Davis, S., J. Haltiwanger, and S. Schuh: 1996, 'Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts'. *Small Business Economics* **8**(4), 297–315.

Donaldson, D. and R. Hornbeck: 2016, 'Railroads and American Economic Growth: A Market Access Approach'. *Quarterly Journal of Economics* **31**(2), 799–858.

Dorn, H.: 1937, 'The Possibility of Underenumeration in the 1920 Census. Part B'. Technical report, National Resources Committee. Population Statistics.

Duranton, G. and D. Puga: 2001, 'Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products'. *American Economic Review* pp. 1454–1477.

Field, A. J.: 2008, 'The Impact of the Second World War on US Productivity Growth'. *Economic History Review* **61**(3), 672–694.

Fishback, P. and J. A. Cullen: 2013, 'Second World War spending and local economic activity in US counties, 1939-58'. *Economic History Review* **66**(4), 975–992.

Florida, R.: 2002, 'The Economic Geography of Talent'. *Annals of the Association of American Geographers* **92**(4), 743–755.

Foster, L., J. Haltiwanger, and C. Syverson: 2008, 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?'. *American Economic Review* **98**(1), 394–425.

Frank, M. W.: 2009, 'Inequality and Growth in the United Sttaes: Evidence from a New State-Level Panel of Income Inequality Measures'. *Economic Inquiry* **47**(1), 55–68.

Glaeser, E.: 2011, *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*. Penguin.

Goldin, C.: 1998, 'America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century'. *Journal of Economic History* **58**(02), 345–374.

Goldin, C.: 2006, 'The Quiet Revolution That Transformed Women's Employment, Education, and Family'. *American Economic Review* **96**(2), 1–21.

Goldin, C. and L. F. Katz: 1999, 'The Returns to Skill in the United States across the Twentieth Century'. National Bureau of Economic Research WP #7126.

Gordon, R. J.: 2016, *The Rise and Fall of American Growth: The US Standard of Living since the Civil War*. Princeton University Press.

Grossman, G. and E. Helpman: 1991, *Innovation and Growth in the World Economy*. MIT Press.

Gutmann, M. P., S. M. Pullum-Pinon, and T. W. Pullum: 2002, 'Three Eras of Young Adult Home Leaving in Twentieth-Century America'. *Journal of Social History* **35**(3), 533–576.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg: 2001, 'The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools'. National Bureau of Economic Research Working Paper # 8498.

Haltiwanger, J.: 2012, 'Job Creation and Firm Dynamics in the United States'. In: *Innovation Policy and the Economy, Volume 12*, NBER Chapters. National Bureau of Economic Research, Inc, pp. 17–38.

Hopenhayn, H.: 1992, 'Entry, Exit, and Firm Dynamics in Long Run Equilibrium'. *Econometrica* **60**, 1127–1150.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow: 2013, 'The Allocation Of Talent And U.S. Economic Growth'. National Bureau of Economic Research Working Paper # 18693.

Hsieh, C.-T. and P. J. Klenow: 2014, 'The Life Cycle of Plants in India and Mexico'. *Quarterly Journal of Economics* **129**(3), 1035–1084.

Hurst, E. and B. W. Pugsley: 2011, 'What Do Small Businesses Do?'. *Brookings Papers on Economic Activity* pp. 73–142. Fall.

Jaffe, A. B., M. Trajtenberg, and R. Henderson: 1993, 'Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations*'. *Quarterly Journal of Economics* **108**(3), 577.

Jaworski, T.: 2015, 'World War II and the Industrialization of the American South'. Working paper, Queen's University.

Jones, B. F.: 2010, 'Age and Great Invention'. *Review of Economics and Statistics* **92**(1), 1–14.

Jones, C. I. and J. Kim: 2014, 'A Schumpeterian Model of Top Income Inequality'. National Bureau of Economic Research Working Paper # 20637.

Jones, C. I. and P. M. Romer: 2010, 'The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital'. *American Economic Journal: Macroeconomics* **2**(1), 224–245.

Jovanovic, B.: 1982, 'Selection and the Evolution of Industry'. *Econometrica* **50**, 649–670.

Kaldor, N.: 1961, *Capital Accumulation and Economic Growth*. Springer.

Kerr, W. R.: 2008, 'Ethnic Scientific Communities and International Technology Diffusion'. *Review of Economics and Statistics* **90**(3), 518–537.

Khan, B. Z.: 2009, *The Democratization of Invention: Patents and Copyrights in American Economic Development, 1790-1920*. Cambridge University Press.

Khan, B. Z. and K. L. Sokoloff: 2004, 'Institutions and Democratic Invention in 19th-Century America: Evidence from" Great Inventors," 1790-1930'. *American Economic Review* **94**(2), 395–401.

King, R. G. and R. Levine: 1993, 'Finance and Growth: Schumpeter Might Be Right'. *Quarterly Journal of Economics* **108**(3), 717–737.

Klein, A.: 2013, 'New State-Level Estimates of Personal Income in the United States, 1880-1910'. In: C. Hanes and S. Wolcott (eds.): *Research in Economic History (Volume 29)*. Emerald Group Publishing Limited, pp. 191–255.

Klette, T. J. and S. Kortum: 2004, 'Innovating Firms and Aggregate Innovation'. *Journal of Political Economy* **112**(5), 986–1018.

Lamoreaux, N. R. and K. L. Sokoloff: 1999, 'Inventive activity and the market for technology in the United States, 1840-1920'. National Bureau of Economic Research WP # 7107.

Lerner, J.: 2002, '150 Years of Patent Protection'. *American Economic Review* **92**(2), 221–225.

Li, G.-C., R. Lai, A. D'Amour, D. M. Doolin, Y. Sun, V. I. Torvik, A. Z. Yu, and L. Fleming: 2014, 'Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975–2010)'. *Research Policy* **43**(6), 941 – 955.

Long, J. and J. Ferrie: 2013, 'Intergenerational Occupational Mobility in Great Britain and the United States since 1850'. *American Economic Review* **103**(4), 1109–1137.

Lucas, R. E.: 2009, 'Ideas and Growth'. *Economica* **76**(301), 1–19.

Lucas, R. E. and B. Moll: 2014, 'Knowledge Growth and the Allocation of Time'. *Journal of Political Economy* **122**(1).

Martin, R. F.: 1939, 'National Income and its Geographic Distribution'. Technical report, Conference Board Studies in Enterprise and Social Progress.

Moser, P., A. Voena, and F. Waldinger: 2014, 'German-Jewish Emigres and US Invention'. *American Economic Review* **104**(10), 3222–55.

Mowery, D. C.: 2010, 'Military R&D and Innovation'. In: *Handbook of the Economics of Innovation*. Elsevier, Chapt. 29, pp. 1219–1256.

National Academy of Sciences: 1964, *Federal Support of Basic Research in Institutions of Higher Learning*. Washington D.C.: National Research Council.

Nelson, R. and G. Wright: 1992, 'The Rise and Fall of American Technological Leadership: The Postwar Era in Historical Perspective'. *Journal of Economic Literature* **30**(4), 1931–64.

Payne, C. S.: 1992, *The Ames Project: Administering Classified Research as a Part of the Manhattan Project*. Iowa: Iowa State University PhD Thesis.

Perez-Gonzalez, F.: 2006, 'Inherited Control and Firm Performance'. *American Economic Review* **96**(5), 1559–1588.

Perla, J. and C. Tonetti: 2014, 'Equilibrium Imitation and Growth'. *Journal of Political Economy* **122**(1), 52–76.

Perlman, E. R.: 2016, 'Dense Enough To Be Brilliant: Patents, Urbanization, and Transportation in Nineteenth Century America'. Working paper, Boston University.

Petro, D.: 2016, 'The 1940 Census: Employment and Income'. *Prologue Magazine* **44**(1).

Philippon, T. and A. Reshef: 2012, 'Wages and Human Capital in the U.S. Finance Industry: 1909-2006'. *Quarterly Journal of Economics* **12**(4), 1551–1609.

Piketty, T. and E. Saez: 2003, 'Income Inequality in the United States, 1913-1998'. *Quarterly Journal of Economics* **118**(1), 1–41.

Romer, P. M.: 1990, 'Endogenous Technological Change'. *Journal of Political Economy* **98**(5), S71–102.

Sarada, M. J. Andrews, and N. L. Ziebarth: 2016, 'The Demographics of Inventors in the Historical United States'. Working paper, University of Iowa.

Schmookler, J.: 1966, *Invention and Economic Growth*. Harvard University Press.

Squicciarini, M. P. and N. Voigtlaender: 2015, 'Human Capital and Industrialization: Evidence from the Age of Enlightenment'. *Quarterly Journal of Economics* **130**(4), 1825–1883.

Stephan, P.: 2014, 'The Endless Frontier: Reaping What Bush Sowed?'. In: *The Changing Frontier: Rethinking Science and Innovation Policy*, NBER Chapters. National Bureau of Economic Research, Inc, pp. 321–366.

Thompson, P.: 2006, 'Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations'. *Review of Economics and Statistics* **88**(2), 383–388.

Online Appendix for "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age"

by Ufuk Akcigit, John Grigsby, Tom Nicholas

June 7, 2017

- Not for Publication Unless Requested -

A Census Data Description

In this section we detail our Census data and the methods we use to prepare it for analysis. We use IPUMS complete-count data from the decennial Censuses in 1880, 1900, 1910, 1920, 1930, and 1940. We are limited to this set of years because the complete Census records are released only with a 72 year lag. In addition, the 1890 Census was largely destroyed in a fire in 1921. We start in 1880 because the Census for this year is systematic and contains a set of family related variables we use in our analysis. As the codebook for the 1880 Census writes:

> *"The 1880 Census is in several critical respects the first "modern" Census; it broke new ground in its completeness of coverage, accuracy of enumeration, and range and detail of questions. The supervision of enumerators shifted from a part-time responsibility of regular U.S. marshals to 150 Census Supervisors specifically appointed for the purpose. To make a full, accurate, and speedy enumeration practical, the size of enumeration subdistricts was reduced from a maximum of 30,000 inhabitants in 1870 to a maximum of 2,500 in 1880 while the number of enumerators was increased from 6,530 to 31,282. A variety of new questions were added that greatly enhance the value of the 1880 Census compared to earlier years. It was the first federal Census to inquire about marital status ... Equally important, a question on relationship to head of family was added, which makes it possible to distinguish kin from secondary individuals and allows construction of a wide variety of variables on family structure."*

The set of variables contained in the Census varies greatly over time. In addition, the micro-data from the 1940 Census is continuing to be populated with additional variables. Table A-1 summarizes the information available in our six decennial Census years.

A.1 Cleaning the Census Data

The Census provides a unique identifier for each individual in its records. These person identifiers, or "PIDs," are unique within Censuses, but are not constant across each Census year: an individual with PID 1 in 1880 is not the same individual with the PID 1 in 1900. We are unable therefore to create a panel dataset using our six Census datasets. Although the PIDs are unique

Table A-1: Variables in the Censuses

Census Year	1880	1900	1910	1920	1930	1940
Age	✓	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓	✓
Gender	✓	✓	✓	✓	✓	✓
Marital status	✓	✓	✓	✓	✓	✓
Years married		✓				
Times married						✓
Birth place	✓	✓	✓	✓	✓	✓
Arrival year (immigrants)		✓	✓	✓	✓	✓
Mother’s birth place	✓	✓	✓	✓	✓	✓
Father’s birth place	✓	✓	✓	✓	✓	✓
Head of household	✓	✓	✓	✓	✓	✓
Family number				✓	✓	✓
Children born		✓				
Children living		✓				
Speak English				✓	✓	✓
Read				✓	✓	✓
Write				✓	✓	✓
Attended school				✓	✓	✓
Highest grade schooling						✓
Own home or rent				✓	✓	✓
Home mortgage				✓	✓	✓
Value of home					✓	✓
Radio					✓	
Occupation	✓			✓	✓	✓
Industry				✓	✓	✓
Class of worker				✓	✓	✓
Income						✓

Notes: This list focuses on those variables we use in our analysis and for which a large number of records have non-missing information. Home ownership variables are populated only for select group of individuals, and cannot be robustly matched to patent data.

in the vast majority of states and years, there are occasions in which the same individual shows up twice in the same year. Supposing data entry errors, we drop these duplicate PIDs.[18]

We take steps to impute missing data where it is easy to do so; for instance, we fill in missing age data by calculating the difference between the observed Census year and the individual’s reported birth year.[19]

Before 1940, many variables are coded in strings rather than as categorical variables. For instance, sex variables can take on values “MALE,” “FEMALE,” “M,” “F,” and additional codes indicating unknown. In many cases, these are easy to categorize into numeric categories. However, in certain instances, additional categorization must be done by hand. For instance, the

[18] One individual in Georgia (PID 559409) has consistently non-sensical data, and is thus dropped from the 1900 Census.

[19] A number of individuals in 1900 have negative ages, or some ages above 130 years old. We drop these individuals from our analysis, supposing data entry errors.

race variable often mixes race and nationality. We therefore must make some assumptions as to what nationality corresponds to which race. For example, we classify those reporting that they are "Asian," "Chinese," "Filipino," "Japanese," "Korean," "Mongolian," or "Siamese" as one category "ASIAN."

There are two additional places where such categorization plays an important role in our analysis. First, the occupation variables contain over twenty thousand unique values in 1880, 1920, and 1930. Many of these unique values are the result of misspellings – such as "FARMR" in place of "FARMER" – or due to differences between British and American English, such as "LABOURER" instead of "LABORER." In order to reduce the dimension of the occupation data, we collapse the raw occupation data into three skill groups – low, medium, or high – and three occupation categories: doctors, lawyers, and farmers. The three skill groups are mutually exclusive, and account for 79.4% of individuals with non-missing occupation data. Doctors and lawyers are all high skill, while farmers can be any skill level so long as they appear to work in farm related activities.

The skill classification proceeds as follows. We first classify individuals into low skill occupations using a string match. Low skill individuals perform routine jobs, sell their labor as hired hands, or work as servants or maids. For instance, if an individual reports an occupation containing the string "ASSIST," "CLERK," "LAUNDR," or "FARM," they are initially classified as low skill.[20] This method will classify those who say they are a shop clerk, blacksmith's assistant, laundry girl, or farm hand as low skill individuals. However, it will also classify farm supervisors and legal clerks as low skill. To correct for this, we next begin the classification of middle skill occupations.

Middle skill individuals are 1) those with particular specialties, such as carpenters or blacksmiths 2) those who perform middle management roles such as supervisors, or foremen, and 3) those in the clergy or law enforcement. We replace those coded with low skill occupations with a middle skill code if the individual both reports a string associated with a medium skill job, and is not an assistant or apprentice. Therefore, those who report that they are a "Foreman on a farm" will initially be classified as low skill because their occupation includes the string "farm," but will be updated to medium skill due to the string "Foreman." On the other hand, a "Blacksmith's apprentice" will not be updated to medium skill, as the string "apprentice" disqualifies classification as middle skill, even though the individual works with a blacksmith.

A similar routine is carried out for high skill classifications. Individuals are classified as high skill if 1) their occupation requires higher cognitive thought, such as a scientist, lawyer, or financier, 2) they are owners, directors, or upper management of ventures, 3) they are highly skilled manual workers, such as jewellers, goldsmiths, or silversmiths, or 4) they are public officials such as members of congress, or politicians. In addition, they must not be assistants, apprentices or hired hands. Once again, therefore, one who "Owns a farm" will initially be a low skill individual, but will be updated to high skill as a result of the string "own." Finally, students and those retired have missing occupation skills.

The occupation categories are more straightforward to classify. Doctors are those who are

[20]This example is far from the full set of strings used to classify individuals. A full list of terms is available from the authors upon request. The set of strings was chosen by hand after examining the most common occupations.

both high skill and who report an occupation string containing "DOCTOR," "MEDIC," "MD," "PHARM," "DENT," "PSYCH," or "OPTOM," among others. Thus pharmacists, dentists, psychiatrists, and optometrists will all be classified as doctors. Lawyers are high skill individuals with an occupation string containing "LAW," "JUDG," "ATTORN," and a number of legislator strings such as "SENATE." Thus attorneys, lawyers, judges, and legislators all count as lawyers by our broad definition. Finally, farmers are any individual who have an occupation string containing broad categories and common misspellings like "FARM," "FRM," "FIELD," and "CROP," or more narrow strings such as "HUSKER," "COTTON," "PICK," or "CHICKEN."

The second major instance in which careful classification is required is in determining the birthplace of individuals. The majority of individuals report their place of birth at the state or country level. However, many give more specific answers such as the city, county, or (if abroad) principality of birth. In order to calculate robust migrant flows, it is necessary to aggregate these more refined answers to a state or country level. While there are too many small cities listed to code each person by hand, we make substantial progress in matching individuals to their state of birth: 86.7% of Census records with non-missing birthplace information are successfully matched.

We begin this refinement process by standardizing place names to be upper case, with no spaces. Next, we assign the largest cities in each state to its logical destination. For instance, "MOBILE" and "BIRMINGHAM" are assigned to Alabama. Note that individuals who were born in the much smaller town of Birmingham, Connecticut, for instance, will be incorrectly matched to an Alabama birthplace. While we are comfortable with this small error in most cases, it can prove quite difficult to address for city names that are large in multiple places. For instance, many people live in both Kansas City, Kansas, and Kansas City, Missouri. In such cases, we assume that the individual did not migrate across state lines if possible. That is, we assign an individual's birthplace to be Kansas if they currently live in Kansas and to be Missouri if they currently live in Missouri. These large cities that appear in multiple states are, as far as we can tell, only cities that straddle state lines. Therefore this conservative approach to migration appears to be justified - even if an individual moves from Kansas City, MO to Kansas City, KS, he will still be living in the same metropolitan area. Since classifying this individual as a migrant is thus misleading, we believe this no-migration error is justified.

A similar routine is carried out for international migrants as well. In particular, many German migrants provided specific states of birth, such as Bavaria, Württemberg, or Hamburg. Again, we aggregate these to the country level. We then divide the reported countries into nine regions: Western Europe, Scandinavia, Eastern Europe, Oceania, Africa, the Middle East, Latin America, Canada, and East Asia.[21]

With the cleaned birthplaces, we can then define the migration status of individuals. An individual is said to be an international migrant if they were born in any of the nine global regions defined above. An individual is defined as an interstate migrant if their birth state is different to their state of residence in the Census. Although we cannot calculate year-on-year migration flows, we can ask whether an individual has moved out of his state of birth, and has

[21] A full list of classifications at both the state and country level is available from the authors upon request.

yet to move back.

A.2 Father Match

In order to study social mobility and the role of parental affluence, we attempt to form a robust link between individuals and their parents. The 1940 Census provides an explicit match between individuals and their spouse and parents, so long as they live in the same household. Using a household identifier and a variable giving an individual's person number in the household (e.g. household head is 1, spouse may be 2, father may be 3, etc.), and the person number of their relevant family, we can explicitly ascertain the PID of the individual's family members.

Before 1940, however, we use our own algorithm to determine the PIDs of individuals' family members. First, we generate a family identifier, given by a unique surname, household ID, and city. We then use the relationship to household head variable recorded by the Census to determine the PID of individuals' family members. Specifically, we first limit ourselves to families with only one household head. Then we consider those individuals who report being either the household head or his/her spouse. We extract the PID of these household heads, and assign them to individuals reporting to be either the son or the daughter of the household head. This builds a crosswalk dataset - every son/daughter of a household head is matched to the PID of his/her mother and father.

To test the validity of this matching routine, we use our algorithm on the 1940 data, checking that we match the correct father to the correct child using the person number variables provided by the Census. Our match is nearly perfect: among individuals for whom the Census provides a person number match, we correctly ascertain the PIDs of an individual's father in over 99.99% of every state's population.

While this high match rate is encouraging, it does not suggest that our algorithm is perfect. Instead, it suggests that we are able to successfully replicate the Census' own algorithm for matching parents to children. The match is still subject to two important caveats. First, we can only match individuals to their parents if they live in the same household, and therefore our match may be more successful for younger children or poorer families. Second, we only match parents if at least one of them is the head of the household. In principle, one could attempt to match other members of the family by considering, for instance, the brother and nephew of the household head. However, matching based on non-nuclear family members introduces additional noise and incorrect matches, particularly in cases in which large families reside in the same household. We therefore avoid these kinds of matches in the construction of our final dataset.

B Patent Data Description

Our analysis is based on three main patent datasets we assembled using a mixture of hand entry and optical character recognition (OCR) techniques based on the original patent documents, and information from existing databases. These data are summarized as follows:

- **Patent Dataset A**. 6,675,311 patents. Consists of close to the universe of patents granted by the USPTO between 1836 and 2004 covering the location of the first named inventor listed on the original patent documents down to the city level.

- **Patent Dataset B**. 60,594 patents. Consists of the universe of patents granted by the USPTO for the years 1880, 1900 and 1910 covering both the name and location of the first named inventor down to the city level.

- **Patent Dataset C**. 5,739,225 patents. Consists of the universe of patents granted by the USPTO including the name of all inventors and assignees from 1920 to 2006.

Note that these datasets contain overlapping and sometimes complementary information—for instance, **Patent Dataset A** contains the location of inventors but not their names whereas **Patent Dataset C** contains names but not locations.

We obtained address information for the first inventor from the original patent documents using OCR and hand entry (**Patent Dataset A**). This dataset contains U.S. patents that were granted to both U.S. citizens and individuals living abroad, but in our analysis we obviously limit ourselves to patents granted to individuals and entities based in the United States. In 1880 94% and in 1940 86% of patents were granted to inventors located in the U.S.

For the years 1880, 1900, and 1910 (**Patent Dataset B**), we extracted the name and address of the first inventor listed on the patent document, under the assumption that this individual was the principal inventor of the art. Single inventors were the norm during this time period. In 1880, 1900 and 1910 approximately 92%, 90% and 91% of patents were granted to a single inventor respectively.

From 1920 through 2006, we retrieved the name of every inventor listed on every patent each year using data supplied to us by the European Patent Office (**Patent Dataset C**). In a parallel work, Akcigit et al. (2017c) created a panel dataset with an inventor identifier through disambiguating the inventor data using the algorithm of Li et al. (2014). We use these data in our analysis of inventor careers in Fact 4.

B.1 Citation Adjustment

Our data includes the number of citations each patent receives from patents granted from September 1947, when the USPTO began to note citation data in a systematic way, to February 2008. Thus, we have the full universe of citations received by patents granted during this time period. Citations start in 1947 because a USPTO Notice was issued on December 19th, 1946, instructing examiners to add citations in the published format of the patent, a practice that was incorporated into the *Manual of Patenting Examining Procedure* (paragraph 1302.12).

For patents granted before 1947, the noted citation count is left censored: a patent granted in 1940 will only have citations from patents granted after 1947, but will not have citations from patents between 1941 and 1946. This artificially deflates the number of citations received by patents before 1947, confounding attempts to use citations as an objective measure of a patent's quality. Furthermore, aggregate citation trends may weaken the link between raw citation counts and patent quality. For instance, if patents granted in 1960 cite an average of 5 prior patents, but those granted in 1990 cite 20 patents, one might expect the average citation received from a 1960 patent to be more indicative of a high quality innovation than a citation received in 1990. We therefore adjust the number of citations received by each patent following the quasi-structural approach laid out in Hall et al. (2001).

This approach relies on two critical assumptions. First, we assume that the citation process is *stationary*. That is, we assume that the evolution of citation shares does not change over time: a patent will on average receive a share $\pi_{k\tau}$ of its citations τ years after it is granted, regardless of the grant year. This allows us to project back our adjustment factors to patents filed before the citation data began in 1947. Second, we assume *proportionality*. That is, we assume that the shape of the citation evolution does not depend on the total number of citations received so that highly cited patents are more highly cited at all lags. This allows the application of the same adjustment factor to every patent in our data granted in a given period and belonging to a given patent class.

The adjustment proceeds as follows. We start with the full patent citation network data, keeping only those patents granted in the United States. Let C_{kst} be the total number of citations to patents in year s and technology category k coming from patents in year t.[22] Further, define P_{ks} to be the total number of citations received by patents granted in year s in technological category k. One can then define $\pi_{kst} = C_{kst}/P_{ks}$ to be the average share of citations received by patents in class k in year s from patents granted in year t We assume that π_{kst} is some multiplicatively separable function of grant year, patent category, and a citation lag. That is, we can write

$$\log[\pi_{kst}] = \alpha_0 + \alpha_s + \alpha_t + \alpha_k + f_k(L) \tag{A-1}$$

for $L = t - s$ the lag between cited and citing patent grant years, and $f_k(\cdot)$ some category-specific function of these lags. For our purposes, we define $f_k(L) = \tilde{\gamma}_{k,L}$. We may then estimate equation A-1 using OLS to recover estimates of $\alpha_0, \alpha_s, \alpha_t, \alpha_k$, and $\tilde{\gamma}_{k,L}$ for each value of s, t, k and L in our data.[23] Taking exponentials of equation A-1 yields

$$C_{kst}/P_{ks} = e^{\alpha_0} e^{\alpha_s} e^{\alpha_t} e^{\alpha_k} e^{\tilde{\gamma}_{k,(t-s)}} \tag{A-2}$$

This formulation allows us to standardize citation counts over time and across categories. Specifically, in order to adjust for patent class, cited year, and citing year effects, we weight each citation from a patent in year t to a patent in class k in year s by $\exp\left(-\hat{\alpha}_k - \hat{\alpha}_s - \hat{\alpha}_t\right)$. Each patent's citation counts are therefore reflective of the patent's quality relative to the average patent in some base year and category.[24]

While this procedure accounts for aggregate differences across patent classes and grant years, it does not yet correct for bias arising from the left truncation of citation records. To build

[22]For the purposes of the adjustment, we use technological categories as defined by the NBER patent data. For a detailed description of these data, see Hall et al. (2001).

[23]It is rare for a patent to receive citations more than 30 years after its initial grant date, and thus we top-code the citation lag L to have a maximum value of 30. That is, we define $L = \min\{t - s, 30\}$.

[24]For our purposes, we choose each patent citation to be relative to a patent in the "Other" category granted in 1975, receiving citations from patents also granted in 1975. Mechanically, this corresponds to setting the omitted categories in estimation of equation A-1 to be $k =$ "Other", $s = t = 1975$.

intuition for the truncation correction, consider an example in which each of the estimated α coefficients were 0: the only bias in our citation data arises from the lag. In that case, the assumptions of proportionality and stationarity suggest a natural adjustment factor for a patent granted L years before the 1947 cutoff. Define $G_k(L)$ to be the CDF of the lag distribution: the share of an average patent's citations received within the first L years after its grant. The adjustment factor is then given by

$$\sigma_{k,L} = \frac{1}{1 - G_k(L)}$$

We would then predict that a patent in category k granted in year $1947 - L$ and receiving c citations from patents granted after 1947 would have received $\sigma_{k,L}c$ citations had the USPTO kept track of citations before 1947.[25]

In order to incorporate the year and category fixed effects into this truncation adjustment framework, one must establish a notion of the CDF of the lag distribution conditional on year and category effects. To do so, we interpret the $\exp(\tilde{\gamma}_{k,L})$'s as weights for each patent in the citation data. For instance, if the estimated $\exp(\tilde{\gamma}_{k,L=2})$ is 2, then an average patent is twice as likely to receive a citation after 1 year than in the year of patent grant, conditional on year and category effects. To construct the CDF of citations by lag conditional on year and class effects, we can sum our estimates of $\exp(\tilde{\gamma}_{k,L})$, normalizing the estimated coefficients so that they sum to 1. This gives us our estimate of $G_k(L)$:

$$\hat{G}_k(L) = \frac{\sum_{l=1}^{L} \exp(\tilde{\gamma}_{k,l})}{\sum_{l=1}^{30} \exp(\tilde{\gamma}_{k,l})} \tag{A-3}$$

We can then calculate our truncation adjustment factor as before[26]

$$\hat{\sigma}_{k,L} = \frac{1}{1 - \hat{G}_k(L)}. \tag{A-4}$$

To summarize, the citation adjustment proceeds in four steps:

1. Estimate equation A-1 using OLS to recover $\alpha_0, \alpha_k, \alpha_t, \alpha_s$ and $\gamma_{k,L}$.

2. For each citation made from a patent p' granted in year t to a patent p in class k granted in year s is weighted by

$$\omega_{k,s,t} = e^{-\alpha_k - \alpha_t - \alpha_s}$$

   Define, for each cited patent p, the year- and category-adjusted citation count c to be the sum of the $\omega_{k,s,t}$ it received.

3. Calculate $\hat{G}_k(L)$ according to equation A-3

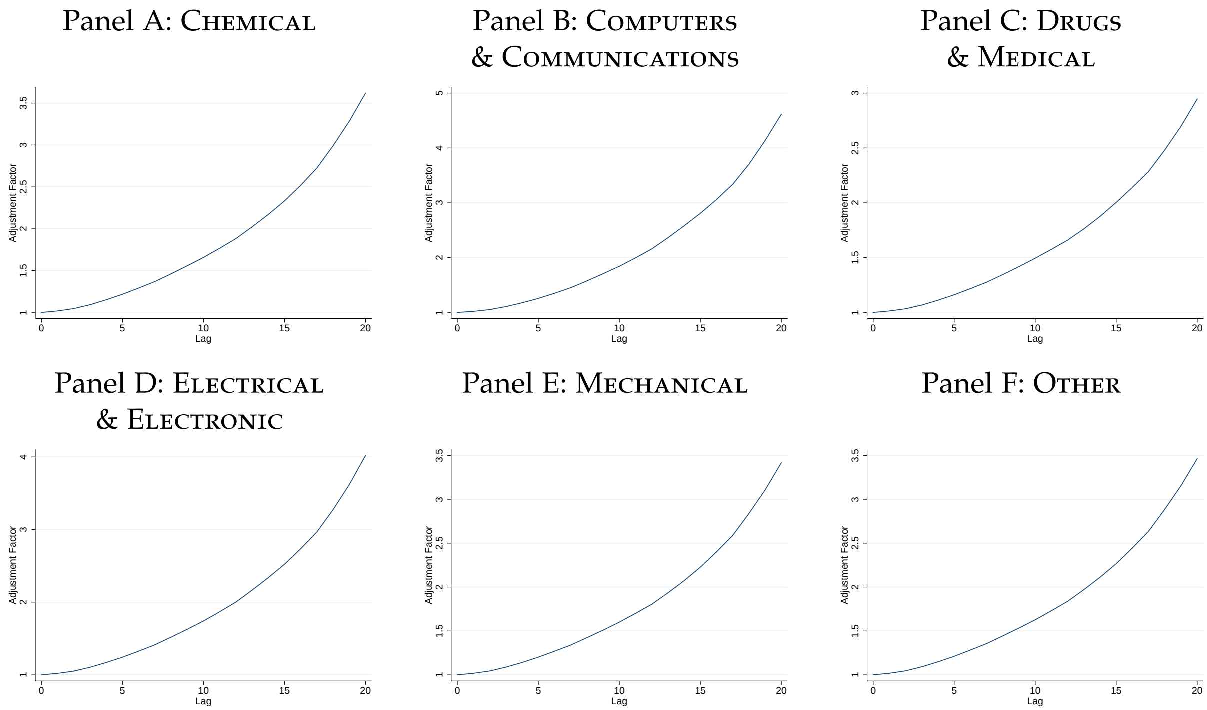
[25] Ignoring year and category effects and adjusting citations in this way does not significantly change the results presented in the main body of the paper.

[26] Note that we only calculate the truncation adjustment up to $L = 20$, despite estimating $\gamma_{k,L}$ for L as large as 30. This is to bound $\hat{G}_k(L)$ away from 1, so that we do not divide by 0 in the adjustment. For L larger than 20, we apply the adjustment factor for $L = 20$.

4. Using $\hat{G}_k(L)$, calculate the truncation adjustment factor $\hat{\sigma}_{k,L}$ according to A-4. Finally, define a patent p's adjusted citation count to be $\tilde{c} = c \cdot \sigma_{k,L}$ if p is in class k and was granted L years before 1947.

Figure A-1 plots the adjustment factors for truncation years for each of the six NBER patent categories. The multiplicative adjustment factors range from 1 to almost 5, and vary by NBER category. Meanwhile, Figure A-2 plots the distribution of log citations and the evolution of the average citation counts according to three adjustment regimes: no adjustment, full adjustment, and an adjustment in which we do not correct for truncation at 1947. We see that the fully adjusted citation counts have a much flatter time series relative to the unadjusted citation counts. This is by design: the purpose of the adjustment is to remove the aggregate fluctuations which do not accurately measure the relative quality of patents.

Figure A-1: Citation Adjustment: Adjustment Factors by Years Truncated



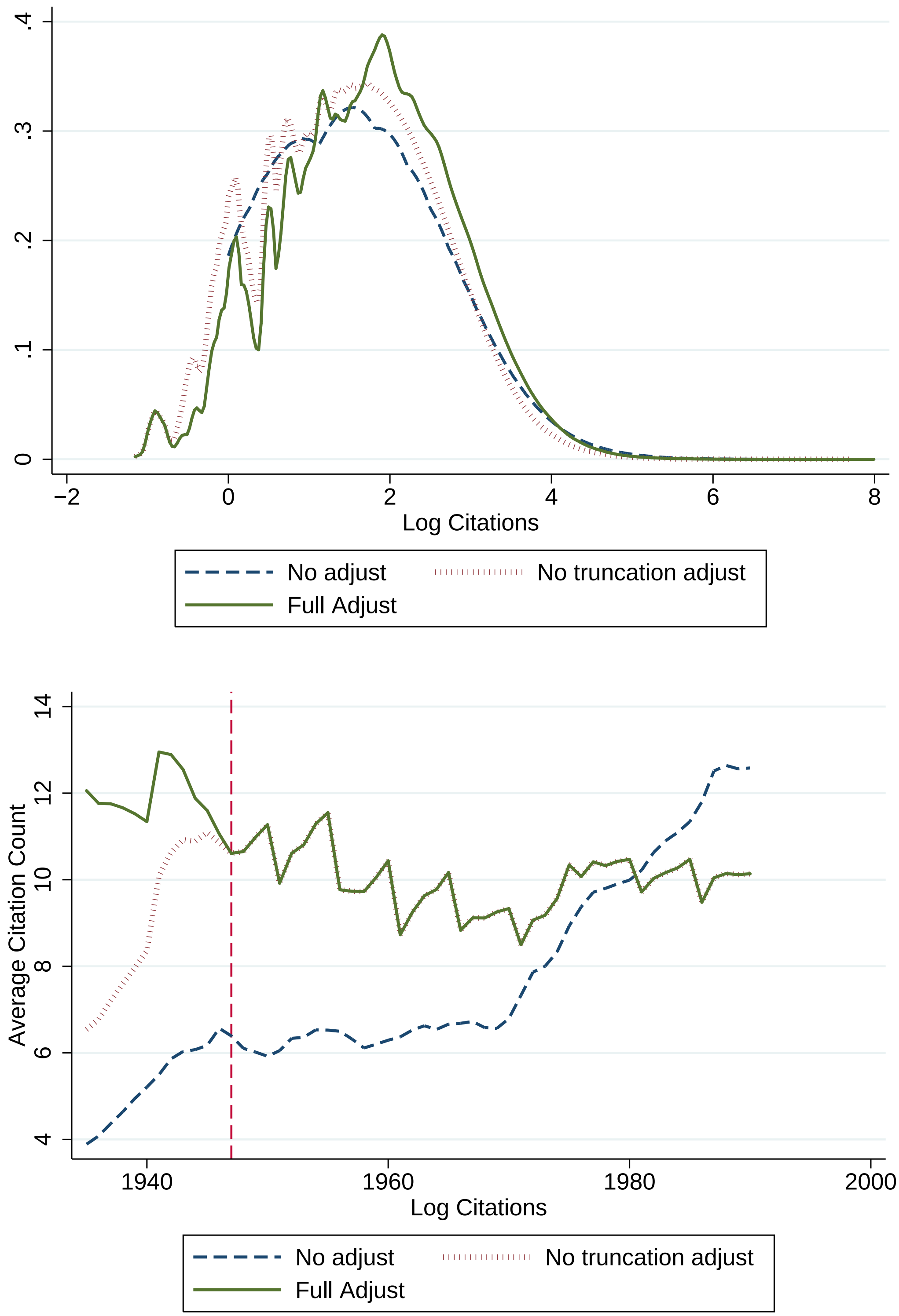
B.2 Patent Classes and Matching Patents to Sectors

We obtain the main USPTO patent class for each patent and the NBER patent aggregations of these classes.[27] We match patents to sectors using the USPTO technology class of the patent.

We also use value added and full-time employment data by sector from 1947 through 1986, before the SIC was revised in 1987. These data are matched to data provided by Bill Kerr containing the fraction of patents in each class which were manufactured and used by every 3-digit SIC code (Kerr, 2008). We first aggregate these SIC codes into the same categories contained in

[27] The USPTO occasionally reclassifies patents based on the emergence of new technologies. Throughout the paper, we use the 2006 classification.

Figure A-2: Citation Adjustment: Citation Distributions and Average Citation Counts over Time



the industry value added data from the BEA. Table A-2 shows this aggregation.

Once we know the fraction of patents in each class that are accounted for by the BEA-provided industries, we assign each class to an industry. We say a patent class c is affiliated with industry

j if industry j manufactures the highest share of patents in class c. We can then calculate the total number of patents for each BEA industry.

C Merging Patent and Census Data

C.1 Data Preparation

We first standardize the names and places listed in the patent and Census data. We begin by ensuring that all names are fully capitalized, and remove all special characters (e.g. ".''s) from the names.[28] In addition, we remove suffixes such as "JR," "Senior," and "III" from listed names. We next parse the names into different words. The surname is taken to be the last word of an individual's name, while an individual's first name is taken to be the first word. The first letter of the second word of an individual's name is taken to be their initial, so long as the name contains at least three words. For example, a name originally recorded as "Thomas Alva Edison," will return three pieces of information: the surname "EDISON," a first name "THOMAS," and an initial "A." Note that this procedure implies that those with multiple words in their surname are constrained to have a first name, single middle initial, and one-word surname. For example, Robert Van de Graaff, inventor of the Van de Graaff generator (a machine that generates static electricity), is eventually listed as "ROBERT V GRAAFF."[29]

Locations are likewise standardized. First, we capitalize all place names listed in the Census and on the patent records. We then ensure that the spelling of common pieces of the place name are constant across the two data sources. For instance, we enforce that the word "SAINT," as in "SAINT LOUIS," are all listed as "ST." In addition, we remove superfluous words such as "WARD," "DISTRICT" or "CITY;" for instance, "NEW YORK CITY" becomes simply "NEW YORK." Finally, we standardize a number of common place names by hand; for example, we impose that the five boroughs of New York City – Brooklyn, Manhattan, Queens, the Bronx, and Staten Island – are all coded as "NEW YORK."

C.2 Merging the Data

We next merge the patent data to the decennial Censuses. To do so, we first insist that records in the Census have the same first name, last name, county, and state as the inventor listed on the patent. In addition, the patent in question must have been granted in the same year as the Census was conducted. While we make a strong effort to clean our data before matching, there remain some cases that do not match even on these basic criteria. Of all patent-inventor instances in the patent data, 70.8% find a match in the census based on these criteria. The remaining 29.2% may not match either because their names were incorrectly entered in either the Census or patent data, or because they may have moved across state lines between the time the Census was conducted

[28] We drop Census records with first or last names longer than 40 characters. We do this because we suppose that such long names arise from input errors.

[29] The 1910 Census provides multiple name fields. We take the most well-populated field, and fill in missing values with the names contained in the other name variables. In the vast majority of cases, the names provided in the two variables are identical.

Table A-2: 2- and 3-Digit SIC Codes and Associated Industry Titles

01-02	Farms	23	Apparel and other textile products	61	Credit agencies other than banks
07-09	Agricultural services, forestry, and fishing	26	Paper and allied products	62	Security and commodity brokers
10	Metal mining	27	Printing and publishing	63	Insurance carriers
11-12	Coal mining	28	Chemicals and allied products	64	Insurance agents, brokers, and service
13	Oil and gas extraction	29	Petroleum and coal products	65-66	Real estate
14	Nonmetallic minerals, except fuels	30	Rubber and miscellaneous plastics products	67	Holding and other investment offices
15-17	Construction	31	Leather and leather products	70	Hotels and other lodging places
24	Lumber and wood products	40	Railroad transportation	72	Personal services
25	Furniture and fixtures	41	Local and interurban passenger transit	73	Business services
32	Stone, clay, and glass products	42	Trucking and warehousing	75	Auto repair, services, and parking
33	Primary metal industries	44	Water transportation	76	Miscellaneous repair services
34	Fabricated metal products	45	Transportation by air	78	Motion pictures
35	Machinery, except electrical	46	Pipelines, except natural gas	79	Amusement and recreation services
36	Electric and electronic equipment	47	Transportation services	80	Health services
371	Motor vehicles and equipment	48	Communications	81	Legal services
372-379	Other transportation equipment	481,482,489	Telephone and telegraph	82	Educational services
38	Instruments and related products	483	Radio and television	83	Social services
39	Miscellaneous manufacturing industries	49	Electric, gas, and sanitary services	86	Membership organizations
20	Food and kindred products	50-51	Wholesale trade	84,89	Miscellaneous professional services
21	Tobacco products	52-59	Retail trade	88	Private households
22	Textile mill products	60	Banking	43,91-97	Government

Notes: Sector codes retrieved from the documentation of value added statistics provided by the BEA: https://www.bea.gov/industry/xls/GDPbyInd_VA_SIC.xls on August 10, 2016.

and the patent was granted. Predictably, this problem is particularly pronounced immediately following the end of the First World War: we match just 61.1% of patent-inventor observations in 1920.

Naturally, there may remain multiple inventor matches if, for example, there are multiple Thomas Edison's living in Middlesex county, New Jersey in 1900. Indeed, 44.4% of our initially matched patents have multiple candidate inventors. We then refine the match further based on other information in the patent documents. First, for each patent, we look to see if one of the candidate inventors in the Census data has the same middle initial as listed on the patent document. If so, we only keep those that match. This removes 8.2% of our multiple matches.

At this stage, the multiplicity concern arises from the possibility of multiple Thomas A Edison's living in Middlesex, NJ in 1900. Thus we refine to a more granular geography. Our second refinement asks whether there are any candidate inventors living in the same city or township as was listed on the patent document. We only keep those who match on this criterion, so long as the patent has at least one matched candidate. Thus we limit ourselves to Thomas A Edison's living in Menlo Park, Middlesex County, New Jersey in 1900. The refinement based on cities removes 7.3% of the duplicate inventors, who survived the refinement based on middle initials.

Multiplicity can still persist, however, and may be particularly common within family units if a son is named after his father. At this stage, both John J Smith Jr and his father John J Smith Sr, living in the same household, would be matched to the same patent. To combat this, we finally refine the match based on an age criterion. For a given patent, we ask if there is a candidate inventor between the ages of 15 and 85 in the Census. If so, we keep that candidate inventor, and discard the candidate children under 15 years old and the elderly above 85. This age refinement removes 5.5% of the multiple inventors present at this stage. We next repeat this refinement with a sharper age criterion, keeping those between 18 and 65 years of age if such a match exists.

Finally, if there are still multiple matches for a given patent, then we exclude the individual and patent from the sample altogether and they are counted neither as an inventor nor as a non-inventor. This is done to be conservative about our match rate, and to avoid inducing spurious correlations from incorrect matches. As a robustness check, we also run our analysis on a sample in which we keep a random inventor for each patent with a multiple match. The results are qualitatively similar, and are available from the authors upon request.[30]

Table A-3 shows the success of our match at each stage of our process. We see that 72.1% of all patents granted in decennial Census years find a match in the Census, while 69.9% of all inventors find a match. Once we impose the county match, these numbers drop to 44.4%

[30]There is one exception to this similarity in headline results. Table 2 in the main body of the paper show a weak correlation between the probability that an individual becomes an inventor and the inventor status of one's father. When we keep a random matched inventor, this correlation becomes large, significant, and positive. This change can be best understood with an example. Suppose that John J Smith Jr is 25 years old and cohabits with his father, the 50 year old John J Smith Sr. The younger John Smith is an inventor of two patents, but his father is not. Because they are both between the age of 15 and 85 with the same first name, last name and middle initial, and live in the same city, we must keep a random John J Smith for each of the two patents. For the first patent, suppose we kept the younger John Smith, while the other patent is assigned to his father. This generates a spurious correlation between an individual's inventor status and that of his father: even though John Smith Sr was never granted a patent, it appears as though he was in our data. These family relationships might be a persistent source of multiplicity, and thus likely drives this particular difference in our results. We therefore favor the more robust results presented in the main body of the paper.

and 45.1%, respectively. Next, we show the percent of all patents that find a unique match at each stage of the process. For instance, 36.6% of patents and 36.5% of inventors have a unique counterpart in the Census data after imposing that Census and patent data match on city.

TABLE A-3: MATCH RATES AT EACH STAGE OF MATCHING PROCESS

	Unit of Observation	
Criteria	Patent	Inventor
Percent with at least one match		
State, Name	72.1%	69.9%
State, County, Name	44.4%	45.1%
Percent with unique match		
State, County, Name	30.3%	29.8%
. . . + Initial	35.1%	35.2%
. . . + City	36.6%	36.5%
. . . + Age between 18 & 65	39.0%	39.2%

We then merge into our data every patent ever granted to each inventor we have successfully matched. Thus, while we only match inventors to the Census if they are granted a patent in a decennial Census year, our matched data contain patents granted to inventors in every year from 1920 through 2006, as well as patents granted to inventors in 1880, 1900, and 1910.

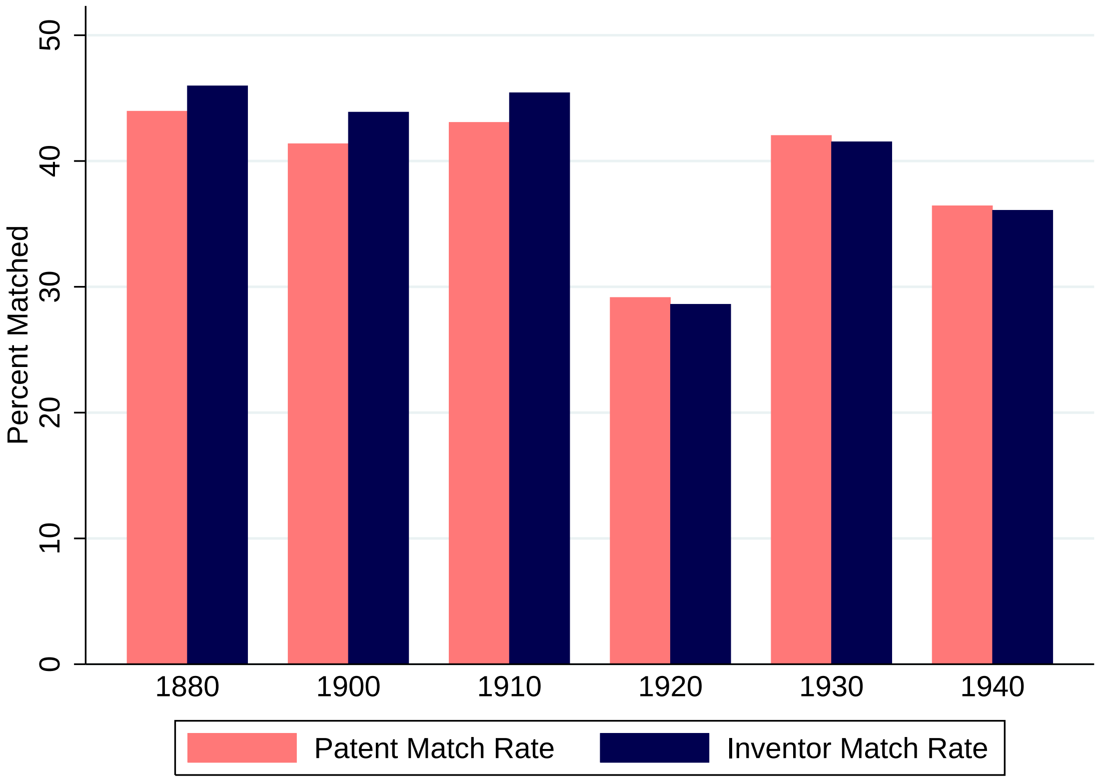
C.3 Match Success

Figure A-3 shows the match rate by decennial Census year. Our most successful match year is 1880, in which we match 44.0% of all patents and 46.0% of all inventors in the patent data to a unique individual in the Census. While the match rate hovers around 40% for most years, in 1920 we match 29.2% of U.S. patents and 28.7% of inventors. The relatively low match rate we observe for 1920 may simply be idiosyncratic. As we point out in Section 2 the 1920 Census was conducted in the winter which led to some level of underenumeration, though not on a scale to bias our results. The effects of World War I demobilization on the movement of ex-servicemen in the population who were also inventors may also have had an effect.

Figure A-4 shows the match rate by state, pooling all years together. Panel A shows the match rate for patents, while Panel B shows the match rate for inventors. There is heterogeneity in the match success across states. While Rhode Island enjoys a successful match rate of 54.5% for patents and 55.9% for inventors, we only match 17.3% of patents and 21.0% of inventors in Nevada. Part of this difference may be attributable to the changing county (and even state) boundaries in the early part of our sample, as frontier states saw rapid increases in population.

A potential concern with our results is that they may be driven by systematic match errors, rather than the unique characteristics for inventors. For instance, if name disambiguation proves especially difficult for common names, our match success will reflect only rare names, which may disproportionately represent a particular race, sex, or age profile. Alternatively, if data input errors are common within the Census, especially uncommon or foreign-sounding names may be matched at a lower rate than traditional American names.

Figure A-3: Match Rate by Decennial Census Year - All States



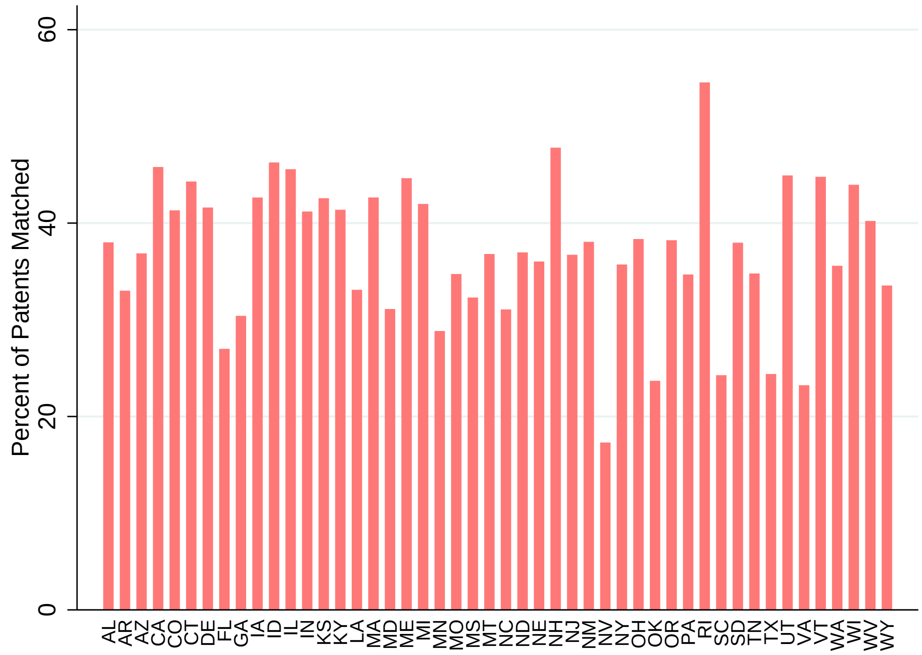
Notes: Figure shows the percent of inventors (solid red bars) and patents (dashed blue bars) present in the patent data who successfully match to the Census data by year. All states are aggregated together to produce this plot.

To test for any biases, we ask whether the patents and inventors that are successfully matched to the Census are observably different from those that are not matched. For this exercise, we consider the universe of patents granted in each of our decennial Census years in the 48 mainland states used in our analysis. We then generate a binary variable equal to 1 if that patent and inventor were successfully matched to our Census data and survived the refinements detailed above. We then regress this indicator on characteristics of the patent and inventor. One might be particularly concerned that we have more success matching common, traditional American names, or particularly prominent inventors. We measure inventor and patent prominence by the number of citations received between 1947 and 2008. We use two indicators for the rarity of the inventor's name. First, we construct the share of the population with each first name using Census data. Second, we include the string length of the inventor's surname.

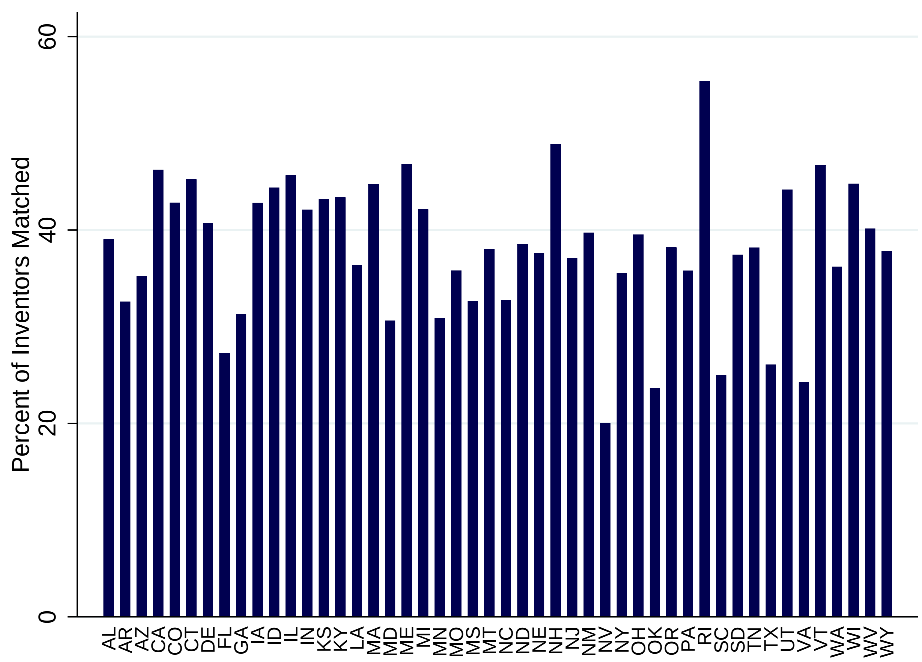
The results of this regression exercise are displayed in Table A-4. Column 1 show that those with longer names are less likely to be matched, and those with common first names are slightly more likely to be matched. However, we do not disproportionately match patents or inventors of a higher quality. These effects are small: increasing name prevalence by 100 (approximately 1 standard deviation) is associated with just a 1.3 percentage point increase in the match rate, roughly 3% of its mean. Meanwhile, a one standard deviation (1.75) increase in an individual's surname length reduces the match rate by 0.6 percentage points.

Figure A-4: Match Rate by State - All Years

Panel A: Patent Match Rate



Panel B: Inventor Match Rate



Notes: Figure shows the percent of inventors (Panel B: dark blue bars) and patents (Panel A: bright red bars) present in the patent data who successfully match to the Census data by the state listed on the patent application. All years are aggregated together to produce this plot.

To test for disproportionate matching of particular population groups, we again use the Census to construct our variables of interest at the first name level. We thus include the percent of

TABLE A-4: SELECTION INTO MATCHING: REGRESSIONS ON PROBABILITY OF MATCH

Panel A: Patent Match Probability

	(1)	(2)	(3)	(4)
Surname Length	-0.346**	-0.219	-0.285*	-0.302**
	(0.154)	(0.154)	(0.153)	(0.149)
Name prevalence (per 10,000 people)	0.013***	0.013***	0.008***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
Citations between 1947-2008	-0.011	-0.010	0.005	-0.004
	(0.011)	(0.011)	(0.007)	(0.007)
Percent First Name Int'l Migrant		-0.203***	-0.201***	-0.201***
		(0.025)	(0.018)	(0.017)
Average Age with First Name		-0.085*	-0.000	-0.012
		(0.051)	(0.036)	(0.036)
Fixed Effects	None	None	State× Year	State× Year Tech Class
Observations	175093	175093	175093	175093
Mean of Dep. Var.	38.65	38.65	38.65	38.65

Panel B: Inventor Match Probability

	(1)	(2)	(3)	(4)
Surname Length	-0.675***	-0.535***	-0.592***	-0.598***
	(0.142)	(0.141)	(0.140)	(0.139)
Name prevalence (per 10,000 people)	0.022***	0.022***	0.015***	0.015***
	(0.002)	(0.003)	(0.002)	(0.002)
Citations between 1947-2008	-0.028***	-0.028***	-0.008	-0.009
	(0.010)	(0.010)	(0.007)	(0.007)
Percent First Name Int'l Migrant		-0.222***	-0.218***	-0.218***
		(0.024)	(0.016)	(0.015)
Average Age with First Name		-0.124***	0.014	0.000
		(0.047)	(0.030)	(0.029)
Fixed Effects	None	None	State× Year	State× Year Tech Class
Observations	122095	122095	122095	122095
Mean of Dep. Var.	39.12	39.12	39.12	39.12

Notes: Dependent variable is an indicator for an observation being matched to the census data, multiplied by 100 for legibility. White heteroskedasticity-robust standard errors reported in parentheses. *, **, and *** represent coefficient statistically different from 0 at the 10, 5, and 1% level respectively. Inventor technology class defined to be the technology class of his/her first patent.

individuals with the inventor's first name who were international migrants and the average age of those with the inventor's name in the Census as dependent variables in columns 2 through 4. Column 2 includes no fixed effects. Column 2 would suggest that, we are less likely to match those with names commonly associated with international migrants, while there is hardly any bias in our age match. This implies that, although we find little difference in the *international* mobility between inventors and non-inventors using our matched data, it remains possible that inventors were more likely to be international migrants. A one standard deviation increase in the percent of people with the inventor's first name who are international migrants (13 percentage points) is associated with a 2.64 percentage point reduction in the patent match rate.

Columns 3 and 4 control for state-year fixed effects in our selection regressions, while column 4 additionally controls for the patent's technology class, and the technology class of the inventor's first granted patent. Column 3 most closely matches that of our previous regression analyses, which use matched data for just one year of the census, and include state fixed effects (see, for example, Table 2). The inclusion of these fixed effects does not significantly change the patterns shown in columns 1 and 2.

Another concern would be that the substantial heterogeneity in state match rates is systematically correlated with key state variables of interest. Although we do not use the matched data for our state-level analysis, it is worth considering this claim. Table A-5 reports estimates from an OLS regression of a state's match rate on its observable characteristics. We see that none of our regional variables predict a state's match rate. Indeed, the full set of variables only explains approximately 12% of the variation in state match rates, as measured by the regression's R^2.

Table A-5: Selection into Matching: State Match Rate Regressions

	Panel A: Patent Match		Panel B: Inventor Match	
	(1)	(2)	(3)	(4)
90-10 Wage Income Ratio	-0.542	-0.364	-0.541	-0.321
	(0.627)	(0.719)	(0.620)	(0.709)
Average Income	-0.003	-0.006	-0.004	-0.007
	(0.012)	(0.017)	(0.012)	(0.017)
Population Density	0.030	0.021	0.034	0.022
	(0.034)	(0.040)	(0.034)	(0.039)
Deposits per capita	0.006	0.005	0.004	0.003
	(0.011)	(0.011)	(0.010)	(0.011)
Average outbound transport cost	0.101	0.134	0.055	0.101
	(0.281)	(0.311)	(0.278)	(0.307)
Percent of residents with college degree	0.083	0.107	0.187	0.209
	(0.721)	(0.773)	(0.713)	(0.762)
Percent employed in manufacturing		-0.132		-0.183
		(0.418)		(0.412)
Percent employed in agriculture		-0.154		-0.196
		(0.242)		(0.239)
Observations	47	47	47	47
R-squared	0.122	0.131	0.114	0.130
Mean of Dep. Var.	37.363	37.363	37.699	37.699

Notes: Dependent variable is the percent of a state's patents matched to the census in one of our six census years. White heteroskedasticity-robust standard errors reported in parentheses. *, **, and *** represent coefficient statistically different from 0 at the 10, 5, and 1% level respectively.

D Additional Robustness Checks

D.1 Sector-level Analysis

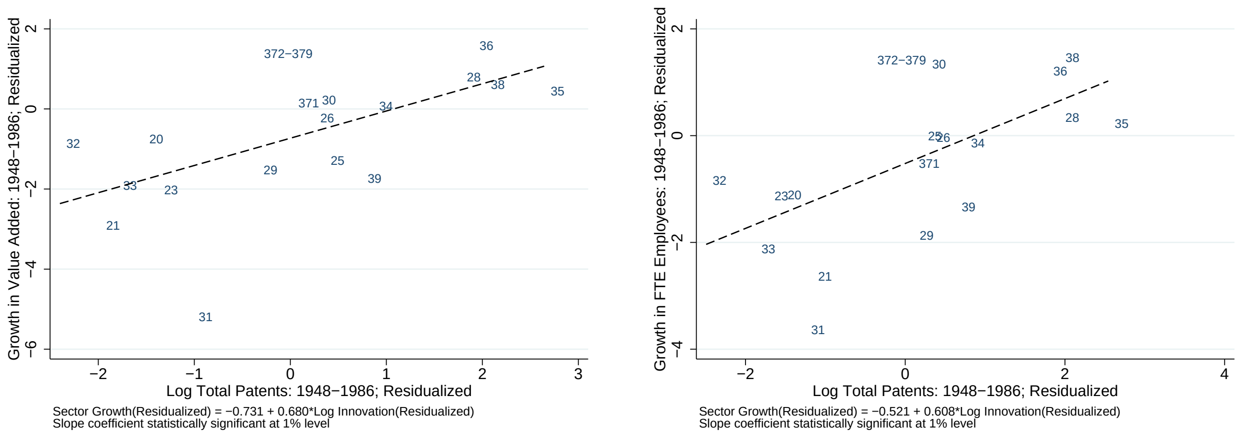
The positive relationship between innovation and output growth persists at the sector-level, as shown in Figure A-5.[31] It plots industry-level annualized growth in value added (Panel A) and

[31] For details on the match between patent classes and industries, see Appendix B.2.

full-time-equivalent employees (Panel B) against the log total patents produced by the industry between 1948-1986, before the change of SIC code definitions in 1987. Both horizontal and vertical axes are residualized against 1948 value added (Panel A) or full-time-equivalent employees (Panel B). Each point represents a 2-digit SIC code.

Figure A-5: Innovation and Long-run Growth: 3-digit Sectors between 1948-1986

Panel A: Growth in Value Added | Panel B: Growth in FTE Employees



Notes: Figure plots industry-level annualized growth in value added (Panel A) and full-time-equivalent employees (Panel B) against the log total patents produced by the industry between 1948-1986, before the change of SIC code definitions in 1987. Both horizontal and vertical axes are residualized against 1948 value added (Panel A) or full-time-equivalent employees (Panel B). Each point represents a 2-digit SIC code, before the codes were changed in 1987. Patent classes are matched to sectors using data provided by William Kerr [3-digit version comes from Kerr (2008) and 4-digit comes from Acemoglu et al. (2016a)]. A patent class k is matched to an industry s if s is the modal user of patents from k. Industry data provided by the Bureau of Economic Analysis.

In Table A-6, we provide the regression coefficients of Figure A-5 and confirm the results using citation-weighted patent counts as our measure for innovation. The results highlight the strong positive association between innovation and economic growth at the sector level.

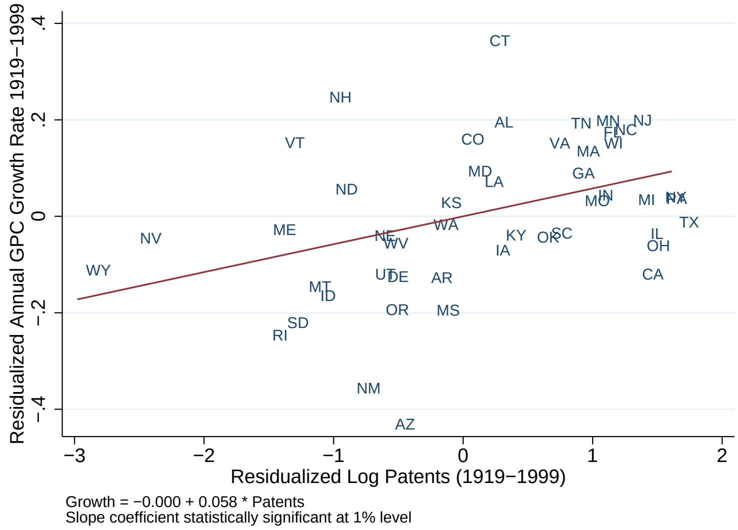
Table A-6: Innovation and Sectoral Growth

Dependent Variable:	Value-Added Growth (1948-1986)		FTE Employee Growth (1948-1986)	
	(1)	(2)	(3)	(4)
Log Patents (1948-1986)	0.679***		0.609***	
	(0.191)		(0.159)	
Log Citations (1948-1986)		0.677***		0.617***
		(0.179)		(0.149)
1948 Dependent Variable Value (1000s)	-0.152	-0.142	-0.627	-0.595
	(0.102)	(0.095)	(0.457)	(0.420)
Observations	18	18	18	18
Mean of Dep. Var.	6.44	6.44	0.39	0.39
S.D. of Dep. Var.	1.61	1.61	1.45	1.45

Notes: Table reports estimated coefficients from a regression in which the dependent variable is the sector-level annualized growth rate in value added (columns 1 and 2) and full-time-equivalent employees (columns 3 and 4). Patent classes are matched to sectors using data provided by William Kerr [3-digit version comes from Kerr (2008) and 4-digit comes from Acemoglu et al. (2016a)]. A patent class k is matched to an industry s if s is the modal user of patents from k. Industry data provided by the Bureau of Economic Analysis. White heteroskedasticity robust standard errors reported in parentheses. *,** ,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level.

D.2 Additional Figures and Tables

Figure A-6: Innovation and Long-run Growth: U.S. States between 1919-1999



Notes: Figure plots the total number of patents granted to inventors in each state between 1919 and 1999 on the horizontal axis, and the annualized growth rate in state GDP per capita between 1919 and 1999 on the vertical axis. Both horizontal and vertical axes plot the variables of interest residualized against 1919 log GDP per capita, to account for conditional convergence. Source: BEA Historical Regional Economic Accounts, and data from (Martin (1939)) courtesy of Price Fishback.

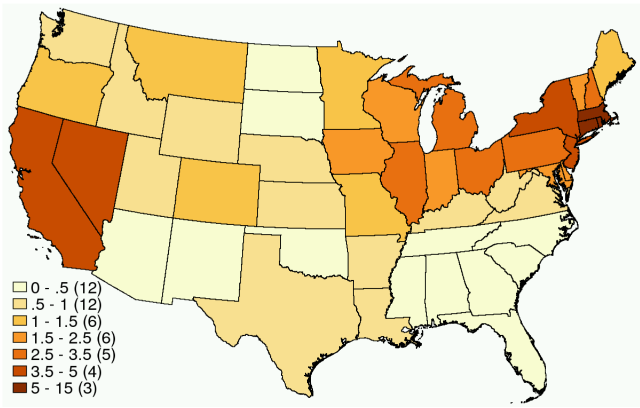
Table A-7: Innovation and Long Run Growth, U.S. States between 1947-1987: Davis, Haltiwanger, and Schuh Growth Rate

	DHS Growth Rate				1^{st} Stage
	OLS	OLS	IV	IV	OLS
	(1)	(2)	(3)	(4)	(5)
Log Patents (1945-1950)	0.039***	0.032***	0.041***	0.027**	
	(0.009)	(0.010)	(0.012)	(0.012)	
OSRD Contracts					0.698***
					(0.083)
Log GDP per Capita (1945)	-0.505***	-0.516***	-0.532***	-0.464***	0.250
	(0.045)	(0.045)	(0.045)	(0.040)	(0.638)
Population Density (1945)		0.332	0.245	0.250	0.574
		(0.198)	(0.173)	(0.177)	(2.291)
1900-1940 GDP DHS Growth Rate				0.114**	
				(0.058)	
Observations	48	48	48	48	48
Mean Growth	0.909	0.909	0.909	0.909	6.698
Std. Dev. of Growth	0.134	0.134	0.134	0.134	1.502
F-Statistic					66.126

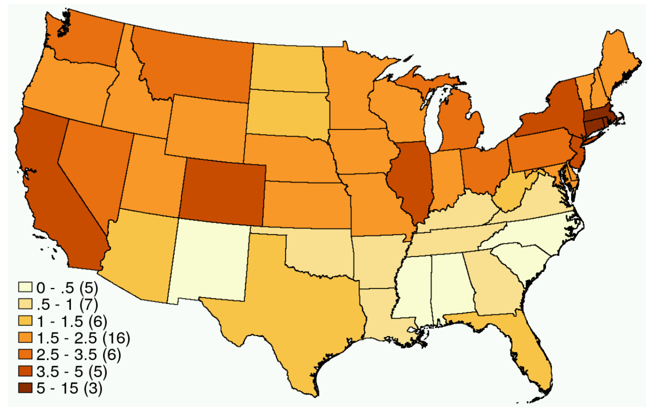
Table reports estimated coefficients from a regression in which the dependent variable is the state-level annualized growth rate in GDP per capita from 1947-1987. White heteroskedasticity robust standard errors reported in parentheses. DHS growth rate refers to the growth rate measure as proposed by Davis, Haltiwanger, and Schuh. The IV estimates are two-stage least squares estimates using the number of OSRD contracts in each state during World War II. *,** ,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level.

Figure A-7: The Geography of Inventiveness over Time: Patents per 10,000

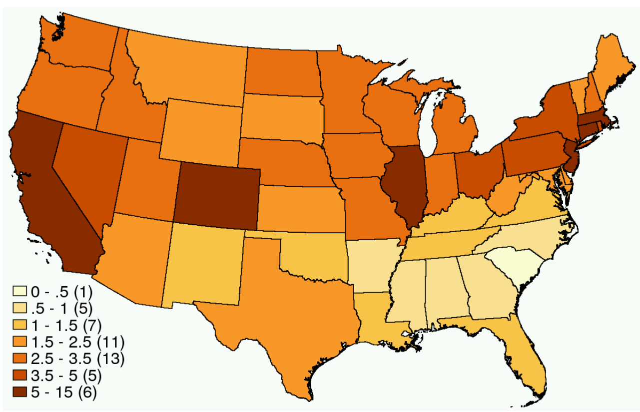
Panel A: 1880



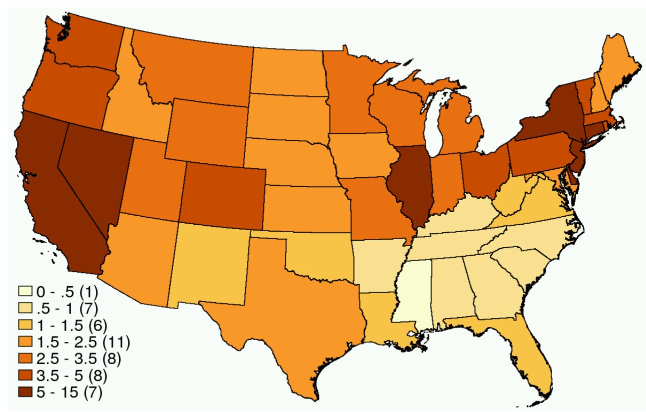
Panel B: 1900



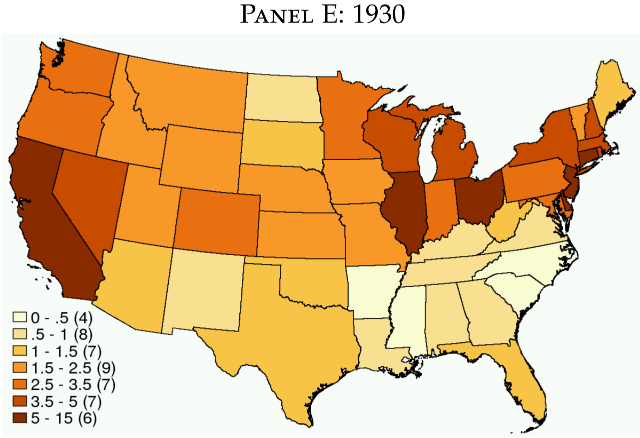
Panel C: 1910



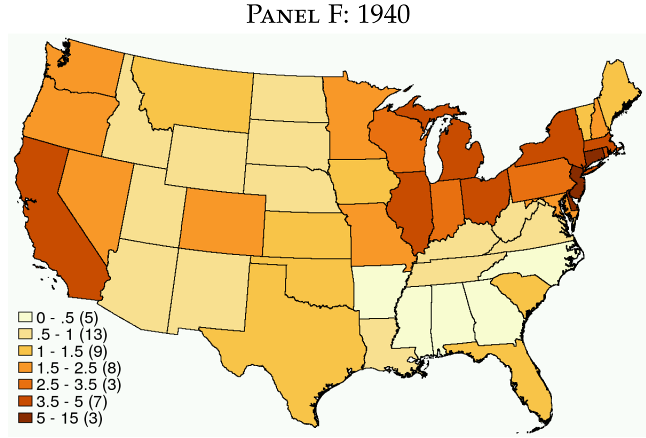
Panel D: 1920



Panel E: 1930



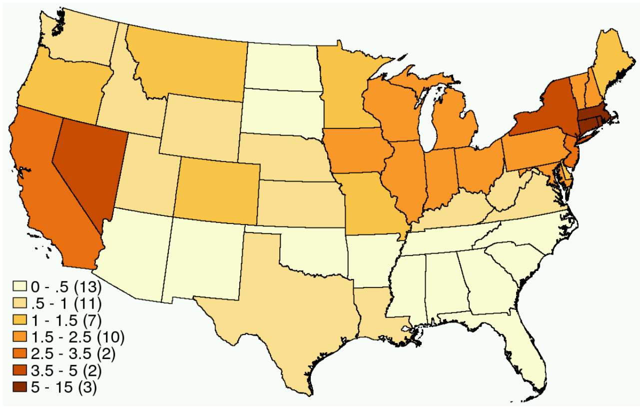
Panel F: 1940



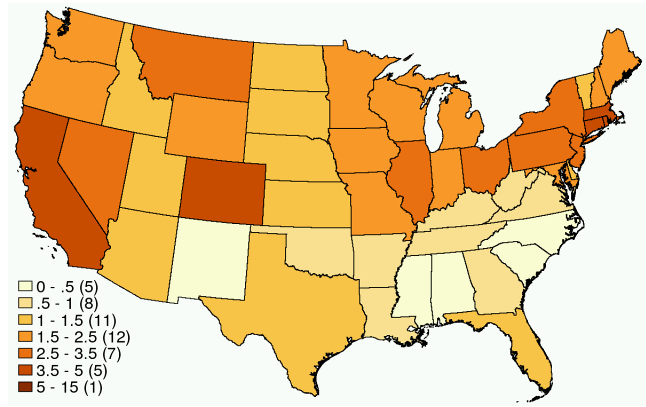
Notes: Figure maps the number of patents per 10,000 residents in each state of the mainland U.S. in each decennial census year of our data. Darker colors represent more inventive activity per resident. Patent data come from the USPTO's historical patent files, while population counts are calculated using the U.S. Census.

Figure A-8: The Geography of Inventiveness over Time: Inventors per 10,000

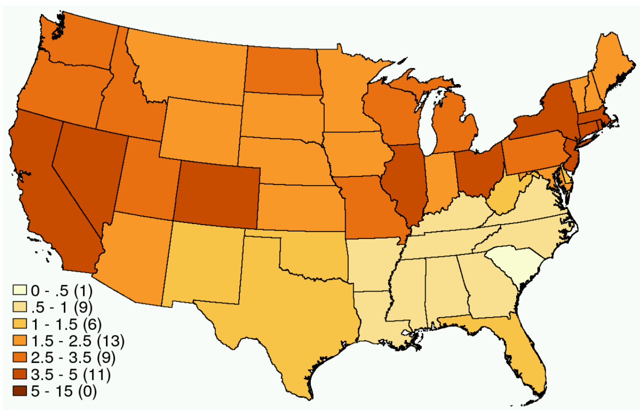
Panel A: 1880



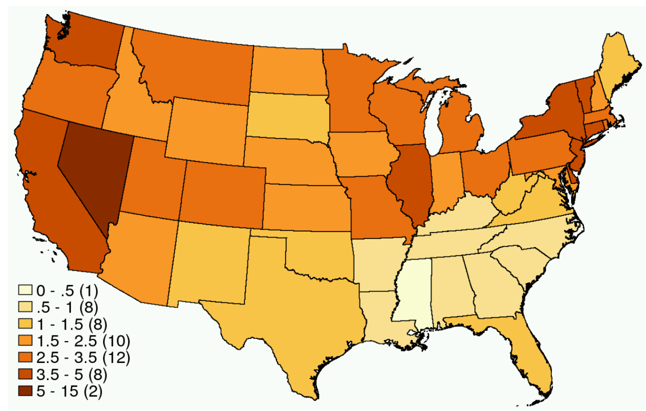
Panel B: 1900



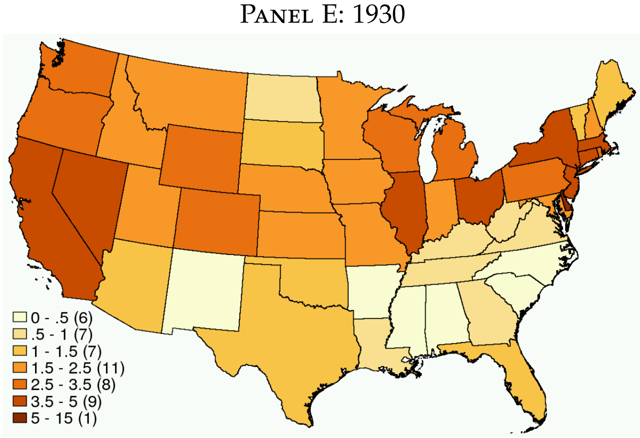
Panel C: 1910



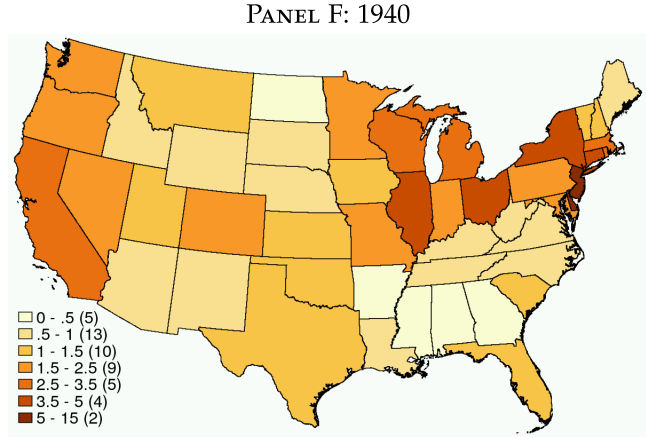
Panel D: 1920



Panel E: 1930



Panel F: 1940



Notes: Figure maps the number of unique inventors per 10,000 residents in each state of the mainland U.S. in each decennial census year of our data. Darker colors represent more inventive activity per resident. Patent data come from the USPTO's historical patent files, while population counts are calculated using the U.S. Census. Source: Historical Census Data, USPTO patent records.

Table A-8: Who Became an Inventor? Controlling for Occupation Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Father Inventor	0.292**	0.292**	0.288**	0.292**	0.289**	0.290**
	(0.135)	(0.135)	(0.135)	(0.135)	(0.135)	(0.135)
Father Income $90^{th} - 95^{th}$ %ile		0.000	-0.003			-0.002
		(0.002)	(0.002)			(0.002)
Father Income 95^{th} %ile and above		-0.001	-0.007***			-0.006***
		(0.002)	(0.002)			(0.002)
Father: High School Graduate				0.002	-0.003	-0.002
				(0.002)	(0.002)	(0.002)
Father: At least Some College				0.001	-0.009***	-0.008***
				(0.002)	(0.002)	(0.002)
Self: High School Graduate			0.008***		0.008***	0.008***
			(0.001)		(0.001)	(0.001)
Self: At least Some College			0.034***		0.035***	0.035***
			(0.005)		(0.005)	(0.005)
Occupation FE	Y	Y	Y	Y	Y	Y
Observations	51078946	51078946	51078946	51078946	51078946	51078946
Mean of Dep. Var.	0.017	0.017	0.017	0.017	0.017	0.017

Notes: Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, a quadratic in age, and father's age. Columns (2) through (5) include indicators for father being between the 50^{th} and 75^{th} percentile of income, and between the 75^{th} and 90^{th} percentile of income as independent variables. The omitted categories are below median income and less than high school eductionn. All columns include fixed effect controls for Census-defined occupation categories, including those with missing occupation data as a separate category. Source: 1940 Historical Census Data, USPTO patent records.

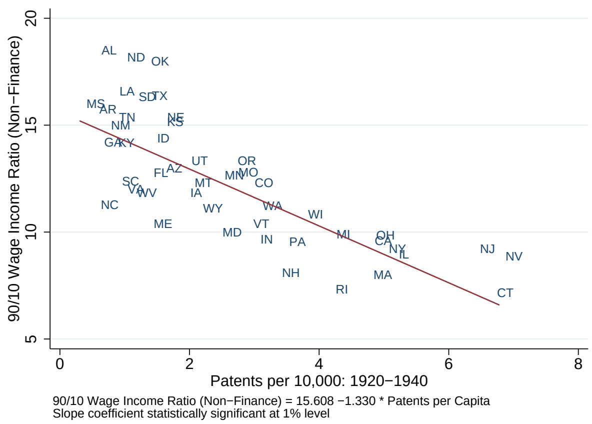
Table A-9: Individual Background and Career Citation Counts

	(1)	(2)	(3)	(4)	(5)
Father Inventor	0.692	0.682	0.544	0.708	0.557
	(0.513)	(0.466)	(0.477)	(0.485)	(0.498)
Father Income $90^{th} - 95^{th}$ %ile		-0.257	-0.244		
		(0.340)	(0.343)		
Father Income 95^{th} %ile and above		0.237	0.170		
		(0.373)	(0.368)		
Father: High School Graduate				0.221	0.118
				(0.220)	(0.220)
Father: At least Some College				0.198	0.073
				(0.175)	(0.175)
Self: High School Graduate			0.006		0.005
			(0.064)		(0.063)
Self: At least Some College			0.280***		0.280***
			(0.048)		(0.049)
Observations	9032	9032	9032	9032	9032
Mean of Dep. Var.	3.205	3.205	3.205	3.205	3.205
S.D. of Dep. Var.	1.964	1.964	1.964	1.964	1.964

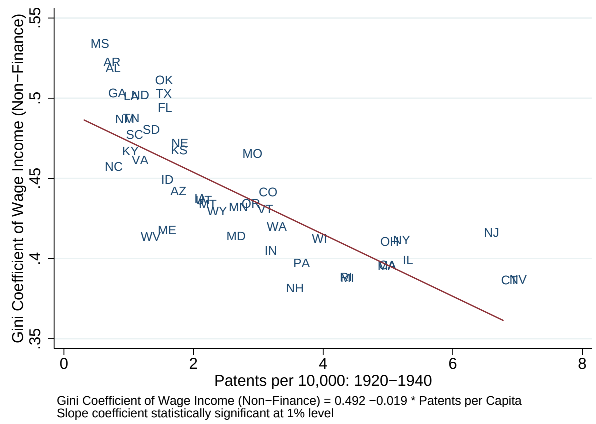
Notes: Table reports coefficients from a regression in which the dependent variable is log one plus career citation counts for the sample of inventors in our matched sample. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, a quadratic in age, and father's age. Columns (2) and (3) include indicators for father being between the 50^{th} and 75^{th} percentile of income, and between the 75^{th} and 90^{th} percentile of income as independent variables. The omitted income category is below median income, and we omit an indicator for the individual having less than a high school education. *,**,*** represent that coefficients statistically differ from 0 at the 10%, 5%, and 1% level. Source: 1940 Historical Census Data, USPTO patent records.

Figure A-9: Relationship between Wage Income Inequality and Inventiveness: Inequality Measures Excluding those Working in Financial Sector

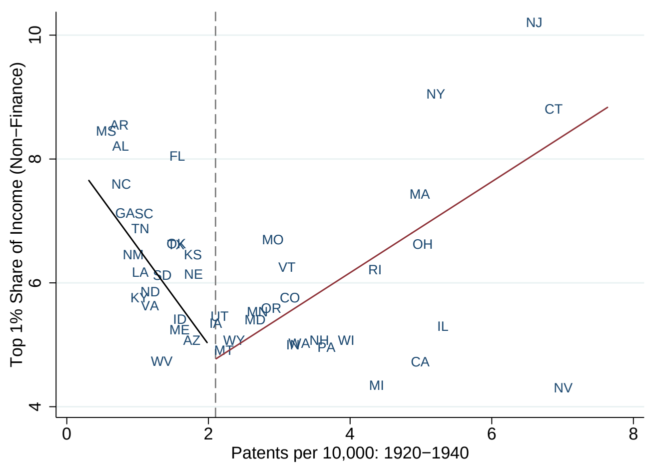
Panel A: Ratio of 90^{th} to 10^{th} Percentile of Income



Panel B: Gini Coefficient

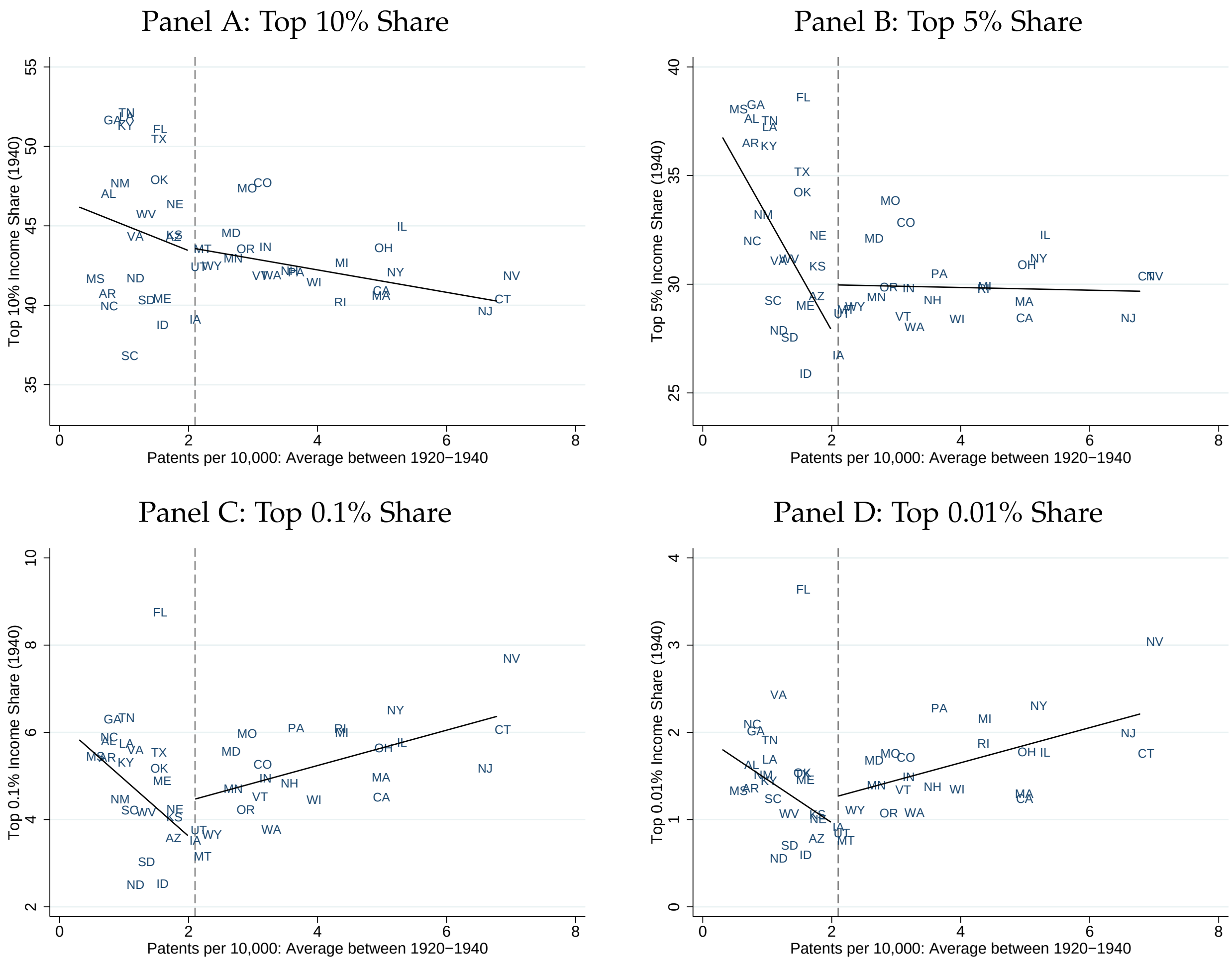


Panel C: Share of Income held by Top 1%



Notes: Figure plots the relationship between average patents per 10,000 residents between 1920 and 1940, and the state-level wage income inequality observed in the 1940 census. All wage inequality measures exclude those who work in the financial sector. Panel A measures income inequality with the ratio of the 90^{th} percentile to the 10^{th} percentile of income, while panel B uses the Gini coefficient as its measure. Panel C measures inequality by the share of income controlled by the top 1% of the state's wage earners. Source: 1940 Historical Census Data, USPTO patent records.

Figure A-10: The Relationship between Innovation and Top Income Shares



Notes: Figure plots the relationship between average patents per 10,000 residents between 1920 and 1940, and the state-level wage income inequality. Source: 1940 Historical Census Data, USPTO patent records, Frank (2009).

Table A-10: T-tests of Difference between Inventor and High-Skill Non-Inventor Wages over the Life Cycle

Age Group	Inventor Mean Log Wage	High-Skill Mean Log Wage	*p*-value
19-25	6.610	6.141	0.000
	(1.025)	(0.922)	
26-35	7.608	6.665	0.000
	(0.703)	(0.857)	
36-45	7.884	6.791	0.000
	(0.725)	(0.906)	
46-55	7.854	6.741	0.000
	(0.793)	(0.942)	
56-65	7.696	6.586	0.000
	(0.977)	(1.000)	

Notes: Table reports average log wages for inventors and high-skill non-inventors within each age group. Wage income data taken from 1940 Census. Standard deviations reported in parentheses below means. Final column presents p-values from a two-sided t-test of means among inventor and high-skill non-inventor populations. Source: 1940 Historical Census Data, USPTO patent records.