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

Ufuk Akcigit John Grigsby Tom Nicholas

*** STILL PRELIMINARY! COMMENTS WELCOME! ***

July 7, 2016

We examine the golden age of American innovation by undertaking a major data matching exercise linking US patent records with complete-count data from Federal Censuses between 1880 and 1940. We identify a casual relationship between patented inventions and long run economic growth and outline a basic framework for analyzing key macro and micro-level determinants. We explore drivers of regional performance including population density, openness to disruption and financial development. We then profile the characteristics of inventors, measure the returns to innovation, and document the relationship between social mobility and invention. Our new data help to address important questions related to innovation and long-run growth dynamics.

**Keywords:** Patents, inventors, demographics, census, earnings, migration.

**JEL classification:** O31, O40, N11, N12.

---

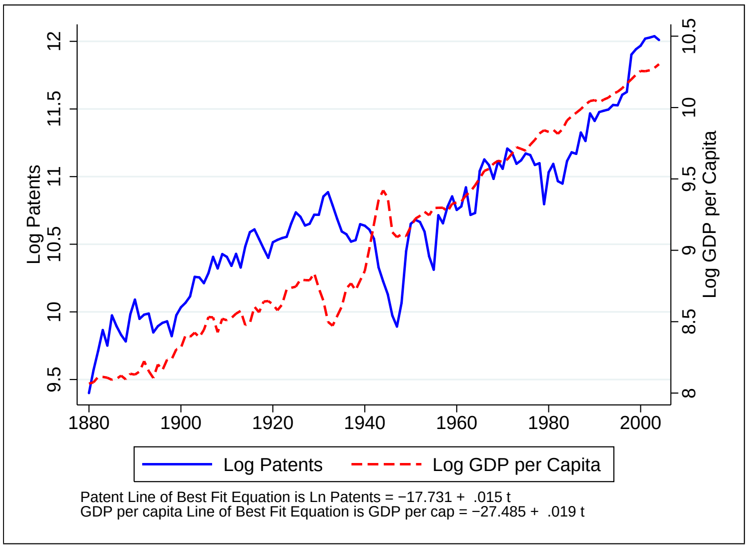
*Authors' Addresses: Akcigit: University of Chicago & NBER (uakcigit@uchicago.edu). Grigsby: University of Chicago (jgrigsby@uchicago.edu). Nicholas: Harvard Business School (tnicholas@hbs.edu). We would like to thank Philippe Aghion, David Galenson, Rick Hornbeck, Ali Hortacsu, Xavier Jaravel, Ben Jones, Pete Klenow, Alan Marco, Casey Mulligan, Rob Shimer, Andrei Shleifer, Stefanie Stantcheva, Nancy Stokey, and Dietrich Vollrath and the seminar and conference participants at College de France, the USPTO and the University of Houston for very helpful feedback and discussions. 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

# 1 Introduction

Innovation and technological progress are engines of long-run economic growth. This basic premise has been at the heart of the 25-year old endogenous growth literature (e.g., Romer (1990), Aghion and Howitt (1992)). Innovations from history, such as light bulbs, air conditioners and storage batteries have tremendous influence today, particularly on the way we live, consume, and produce (e.g., Gordon (2016)). Learning from the creators of these inventions is key to shedding light on several current debates in the innovation and growth literatures. Who were the creators of these essential technologies? What demographic and economic backgrounds gave rise to these innovations? How did these inventors and their innovations influence economic inequality, social mobility and economic growth? To answer these important questions, one needs a large-scale micro-level historical dataset on inventors that can be mapped to both socio-economic variables on individuals, as well as to regional economic aggregates. This paper is a major undertaking in this direction.[1]

For the first time, we provide a systematic match between 220,000 inventors who patented their inventions at the United States Patent and Trademark Office (USPTO) from 1880 to 1940 and 640 million individual records from the US Census Bureau. This major data matching exercise allows us to uncover fundamental facts about a significant number of innovations and their creators during a critical era of US economic progress.

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



Patent Line of Best Fit Equation is Ln Patents = −17.731 + .015 t
GDP per capita Line of Best Fit Equation is GDP per cap = −27.485 + .019 t

Source: USPTO, Maddison, Bureau of Economic Analysis, Klein (2013).

As an overview of the underlying innovation data, Figure 1 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, allowing us to base our analysis on a substantial number of observations. The time period we cover is central to recent debates on innovation and growth. Indeed, we analyze the years

---

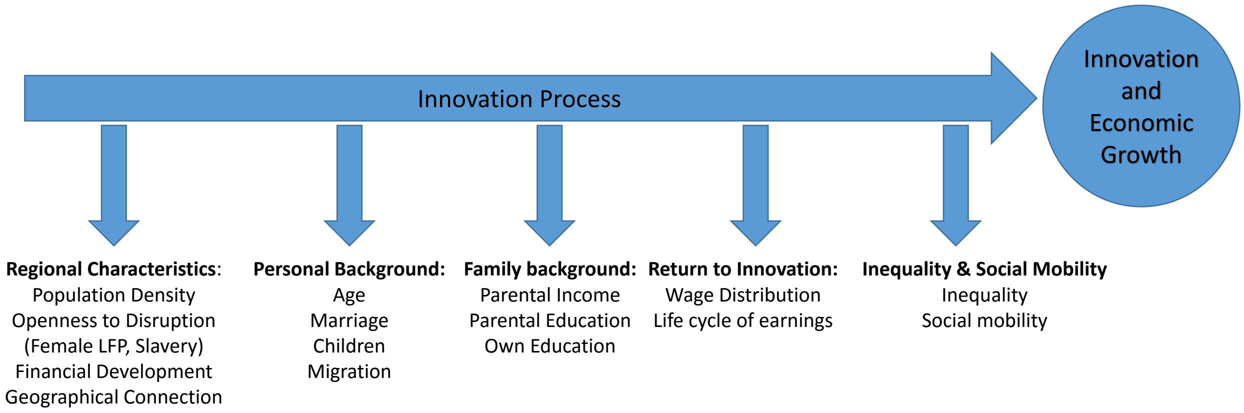
[1] Typically such data has only been available historically for specific samples of superstar inventors (e.g., Khan and Sokoloff (2004)) or for broader populations in modern time periods (Aghion et al. (2015b), Bell et al. (2015)).

that Gordon (2016) associates with the second industrial revolution, which produced major innovations like electricity and the motor vehicle. We also span the 1930s, which Field (2003) identifies as the most innovative decade of the twentieth century. More generally, we find a positive association between our innovation measures and output growth over the long run, in keeping with the predictions of the large theoretical literature highlighting the central role of technological progress in endogenous growth (e.g., Romer (1990), Aghion and Howitt (1992), Aghion et al. (2014)).

**Roadmap** We analyze the innovation process through the experiences of inventors over their life cycles. Some famous case studies can be useful to motivate our approach. Born to a poor family in Ohio, Thomas A. Edison faced tight financing constraints. A creative inventor himself, he wanted to develop the first dedicated research laboratory where inventors could interact and produce new technological developments. For this purpose, he moved to New Jersey where he accessed capital from a group of financiers, including J.P. Morgan, and built the Menlo Park Lab. Edison's experience suggests the importance of funding, population density and human interactions in the innovation process. As another example, Melvin De Groote, one of the most prolific inventors in US history, received two degrees in Chemical Engineering. His highly educated background stands in contrast to that of Nikola Tesla, who dropped out of college. Tesla, a Serbian immigrant, believed that personal relationships detracted from productive research time. Consequently he never married. Tesla's decision shows that inventors faced trade-offs when it comes to time allocation.

The above case studies hint at myriad factors which might spur innovation, each of which we test empirically. To organize our exposition, we provide a framework to consider the innovation process, outlined in Figure 2 below.

Figure 2: ROADMAP OF THE ANALYSIS



First, we exhibit the relationship between innovation and long-run economic growth, as postulated by the active endogenous growth literature. We study the relationship between innovation and growth across states over 100 years between 1900 and 2000. We show that more innovative states and sectors grew faster on average with a doubling of innovative activity in a state being associated with 15% higher GDP per capita after 100 years. We verify that this positive relationship holds in the data at the sectoral level. We also exploit an historical episode of a shift in innovation activity during World War II to show that these estimates are causal.

Our results show that the link between innovation and economic growth has been strongly positive and economically very sizable. Estimates suggest that if two states had the same initial GDP per capita in the beginning of the period and one state innovated four times more than the other (Massachusetts vs Wyoming, for instance), this could lead to 30% higher GDP per capita in the innovative state after 100 years. Hence, we frame our analysis with robust empirical evidence indicating that innovation is a fundamental cause of long-run economic growth.

Next, we study the characteristics of innovative states to explore the macroeconomic environments where inventors flourished. We find that the innovative regions were more densely populated, had higher female labor force participation rates, had lower slave ownership before the Civil War, were more financially developed, and were better connected to other parts of the country by transport links. In particular we argue that the social and economic composition of innovative regions indicates more "openness" to technological disruption (Florida (2002), Acemoglu et al. (2014)) than less innovative regions.

We then transition from the macreconomic environment to the micro-level to determine the demographic and family backgrounds of the inventors of the golden age. Inventors were most productive between ages 35-55, they delayed marriage and had fewer children. Inventors were more likely to have migrated from their state or country of birth (especially after age 35), and moved to more densely-populated, better financially-developed and more open-to-change regions. Inventors had highly-educated parents, were highly educated themselves, and has high-income, even after controlling for observables. Overall, we find strong inventor life-cycle and human capital effects.

Finally, the richness of our data allows us to investigate the relationship between income and inventiveness at both the inventor and regional levels providing new insights into the mobility of American inventors. Innovation is a process where inventors invest in costly effort *ex-ante* in the expectation of gaining *ex-post* returns. We find that inventors had 3 times higher labor income on average, 73% of the inventors were in the top-10% of the overall income distribution and that inventors had a steeper earnings profile over their life cycle. We identify strong returns to the quality of innovation, as inventors with higher citation-adjusted patent counts received higher wage income.

To study regional income dynamics we focus on various measures of inequality: the 90/10 ratio, the Gini coefficient and the top-1% income share. Our preferred social mobility measure focuses on the fraction of those with a low-skill father who themselves have a high-skill occupation. We find that innovative regions 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 we find that the most innovative states had higher social mobility.

Using novel historical microdata, our study provides key macro and micro-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, Section 3 presents the empirical results, and Section 4 concludes.

# 2 Data

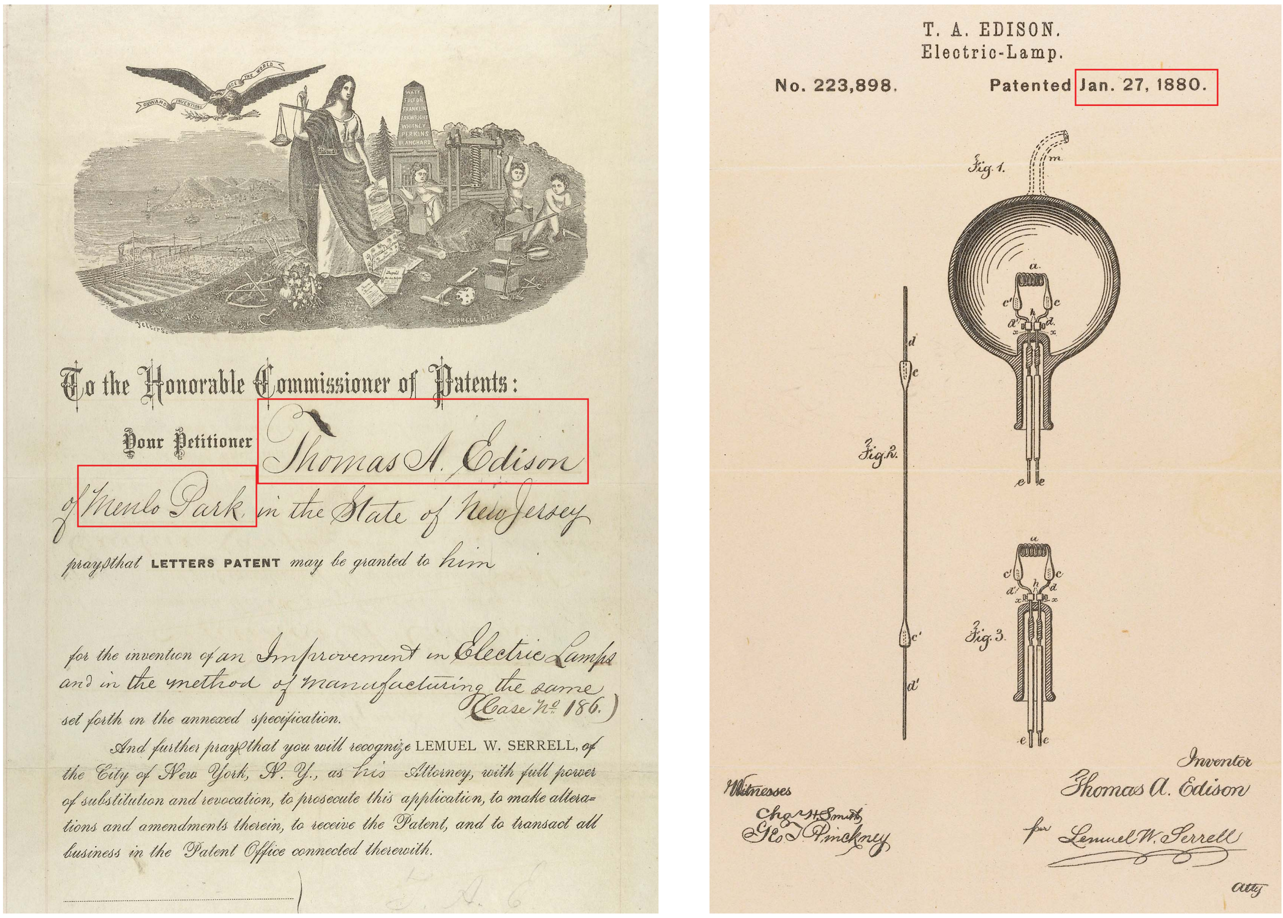
## Census and Patents

The release of the complete-count Census data by IPUMS provides an opportunity to examine a number of questions related to the historical development of innovation in the US. We have conducted a match of the Census data with patent data to establish a profile of inventors active from the 1880s to the 1940s. The main challenge in this exercise is to match inventors, as listed on patent documents, to individuals listed in the Census, in the absence of a unique identifier.[2]

We begin with the hard copies of the historical patent records, such as the famous USPTO patent #223,898 granted to Thomas Edison in Menlo Park on January 27, 1880 (see Figure 3). These records include the name and location of the inventor and the year of the invention as highlighted in red boxes in Figure 3. Through a mixture of OCR and hand-entry we extracted all information on the names of inventors and their locations from USPTO patent documents in years corresponding to the availability of US Census data—1880, 1900 and 1910.

Additionally, for the period 1920 to 2006 we used the same techniques to identify names and locations of inventors for all years. Because our interest is in those individuals who were also in the US Census, we excluded all foreign-domiciled inventors. In 1880 94 percent and in 1940 86 percent of USPTO patents were granted to inventors located in the US.

Figure 3: Front and Back Pages of USPTO Patent # 223,898



To the Honorable Commissioner of Patents:

Your Petitioner Thomas A. Edison of Menlo Park in the State of New Jersey prays that LETTERS PATENT may be granted to him for the invention of an Improvement in Electric Lamps and in the method of manufacturing the same (Case No. 186.) set forth in the annexed specification.

And further prays that you will recognize LEMUEL W. SERRELL, of the City of New York, N. Y., as his Attorney, with full power of substitution and revocation, to prosecute this application, to make alterations and amendments therein, to receive the Patent, and to transact all business in the Patent Office connected therewith.

T. A. EDISON.
Electric-Lamp.

No. 223,898. Patented Jan. 27, 1880.

Witnesses
Chas H Smith
Geo T Pinckney

Inventor
Thomas A. Edison
for Lemuel W. Serrell
atty

---

[2] The first Social Security number was issued in 1936. Unfortunately, no Social Security numbers are included in the 1940 Census, and in any case these were not listed on patent documents at the time.

Our data collection effort for the early part of our study was confined to only the years in which the Census was undertaken. We collected data on patents as of their grant year—the year when the USPTO issued a patent to the applicant. For the early years of our study the correspondence between the timing of a patent being filed by an applicant and the timing of a patent grant was quite close—for example in 1880 an average of 170 days elapsed. By 1930, however, the average pendency period had extended to over 1,000 days and was still over 800 days in 1940. With annual data for post-1920 patents we can be flexible with our matching to address differences in locations given at the time of a patent grant date and in the Census survey years.

A patent entry shows the name and address of an inventor: it shows the surname, first name, middle initial(s) where relevant, state, city, county, and country of the applicant when the patent was granted. It also shows the individual, or more commonly the firm, to whom the patent was assigned, where assignment represented a transfer of ownership rights. It is important to note that even by 1930 over half of US patents were unassigned as of their grant date.

Although patents could be sold because a market for technology had flourished since the middle of the nineteenth century (e.g., Lamoreaux and Sokoloff (1999), Akcigit et al. (2016a)) we can be reasonably sure that the location of the inventor corresponds closely to the location of the invention. We also estimate any potential source of location bias by tracking inventors who can be observed in different places in the patent and the Census data over time.

We obtained complete-count Census data from IPUMS at the standard decennial intervals for years up to 1940. These data represent a substantial improvement over the more restricted samples that had traditionally been available. A Census entry shows (among other things) the name and address of each individual surveyed. We start our matching exercise in 1880 because that is the first year a reasonable number of patent observations become available (around 13,000 patents). The year 1890 is excluded from our analysis because some of the Census schedules were destroyed in a fire in 1921 and others were destroyed under intransigent Federal record management polices existing at the time. Only a limited set of 1890 Census schedules survived.

Beyond the name and location of individuals, the information contained in the Census 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, in other instances variables are recorded in one year only to be dropped in another. Occupation is listed in 1880 but not in 1900 or 1910. We limit ourselves to the working age (18-65) population in the continental United States, as Alaska and Hawaii were not incorporated into the Union until 1912 and 1898, respectively.

Generally, the data becomes more comprehensive in later years. Beginning in 1920 a much wider array of variables are available, specifically relating to education (school attendance) and home ownership. Furthermore, the 1940 Census explicitly 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 (e.g., Abramitzky et al. (2014)).

### Approach to Matching Records

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. For example, in their study of intergenerational occupational mobility, Long and Ferrie (2013) tracked fathers and sons across the 1851 Census and the 1881 Census using the proximity of the name and birth year. In a recent paper, Feigenbaum (2015) estimates intergenerational income mobility using a machine-learning approach to match individuals from the 1915 Iowa State Census and the 1940 US Federal Census. He matches on first name, surname, middle initial, state of birth, and year of birth finding approximately a 59 percent match rate.

Unlike these prior studies we do not observe year of birth in the patent records, 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 patent has the same first name and surname as the individual in the Census, and lives in the same state. Naturally, this leads to a large number of repeated individuals. For some states, we even match up to 20 people for an average inventor in the patent data.

Therefore, second we 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. Next, we ask if there is any matched inventor between 15 and 85 years old. If so, we keep that inventor only. 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—overall our matching rates are encouraging. 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.

We also used variants of this methodology to identify parent-to-children traces across the decennial Census years. This is important for our purposes as we are interested in the mobility of US inventors and hence we want to measure the extent to which innovation was an effective social elevator. In this instance our data has some advantages over that used for perspectives on invention and social mobility in the modern era. Whereas Bell et al. (2015) have identified the parents of children who become inventors, their data are constrained by inventors whose birth year falls after 1980 and who patented by 2012—a thirty two year window. For us, sixty years of patent and Census record data means that more complete changes in mobility levels become observable.

### Technology Areas and Citations

We use two datasets to augment the information available from the original patent documents. First, we use the USPTOs classification of patents to isolate the technology area of inventions. This classification is consistent over time because whenever a new classification is introduced, existing patents are retroactively re-classified. Patents list several technological components, and we observe main classes and subclasses for each invention. Second, we use historical patent citations to identify the most influential technological development. Our data include 3.8 million citations to patents granted between 1880 and 1940 in the population of patents granted between February 1947 (when front page citations began to be systematically recorded) and September 2008.

### Additional Data Sources

Our analysis uses other important datasets. Intercensal population estimates are provided by the Census Bureau. State output data is taken from the Bureau of Economic Analysis (BEA) for 1929 through the present day. Klein (2013) provides estimates of gross state products in 1880, 1890, 1900, and 1910. Sector value added and full-time equivalent employment data come from the BEA.[3]

To establish proxies for financial development, we use data provided by the Federal Deposit Insurance Corporation (FDIC). The FDIC dataset, downloaded from the University of Michigan's ICPSR repository (number 0007) provides the number of deposits, banks, and bank suspensions at a county level from 1920-1936. Transport cost data were obtained from Donaldson and Hornbeck (2016). In our instrumented growth regressions (detailed below) we use data obtained from the Library of Congress on Office of Scientific Research and Development (OSRD) contracts for technological development efforts during World War II.

## 2.1 Summary Statistics

As precursors to the main part of our analysis we present descriptive statistics on our data, which we discuss in more detail in subsequent sections. In keeping with our approach to examining US innovation from macroeconomic and microeconomic perspectives, we structure these data to characterize inventiveness at the state and individual inventor levels.

In Table 1, innovation is measured as the average number of patents per capita between 1880 and 1940 and we distinguish between the most and least innovative states. In Table 2, we present a statistical profile of inventors based on the availability of information in the Census and we compare inventor characteristics with those of the population as a whole.

Finally, as Figure 4 also highlights, it is worth noting at the outset that our data consists overwhelmingly of white male inventors. Our time period roughly coincides with phases Goldin (2006) identifies (i.e., from the 1880s to the 1950s) where women's involvement in the labor market was generally restricted to positions like office and clerical work. Khan and Sokoloff (2004) found only

---

[3] Downloaded from http://www.bea.gov/industry/gdpbyind_data.htm. We use data between 1948 and 1986, as the SIC codes change then. The match from USPTO classes to SIC codes is done using files provided by Bill Kerr and assigns patent classes to the SIC code which manufactures the highest share of patents in that class.

one female inventor in their list of 400 superstar US inventors listed in the *Dictionary of American Biography* who were born before 1886.

Table 1: THE CHARACTERISTICS OF INVENTIVE STATES

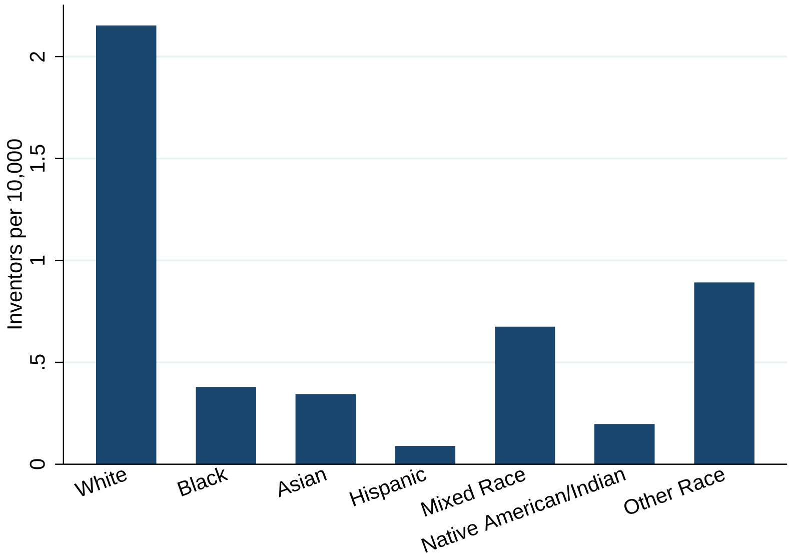
| | Top 10 Inventive States | Bottom 10 Inventive States |
|---|---|---|
| Av. Population (000s) | 2716 | 1437 |
| Population Density (Pop per km$^2$) | 38.85 | 11.35 |
| GDP per Capita | 623.5 | 238.6 |
| Av. Patents per 10,000 People | 5.8 | 0.7 |
| Av. Patents Granted | 1571 | 96.94 |
| Av. Inventors per 10,000 People | 3.33 | 0.57 |
| % Interstate Migrant | 39.80% | 19.96% |
| % International Migrant | 20.64% | 1.68% |
| Migrant Inventors per 10,000 Interstate Migrants | 29.26 | 6.01 |
| Migrant Inventors per 10,000 International Migrants | 22.10 | 7.78 |
| Percent White | 96.2% | 66.8% |
| Percent Black | 2.8% | 30.5% |
| % Over 35 Years Old | 38.5% | 27.7% |
| % Under 35 Years Old | 61.5% | 72.3% |
| No Schooling | 2.88% | 3.95% |
| Less Than High School | 65.16% | 74.69% |
| High School | 20.77% | 12.32% |
| Some College | 6.25% | 5.79% |
| 4+ Years College | 4.94% | 3.24% |
| High Skill Occupation | 9.3% | 4.7% |
| Medium Skill Occ. | 27.4% | 19.3% |
| Low Skill Occ. | 63.3% | 75.9% |
| Employment Rate | 54.69% | 54.33% |
| Mean Wage Income | 1140.7 | 681.4 |
| Median Wage Income | 947.6 | 459.4 |
| Male Labor Force Participation | 89.05% | 88.21% |
| Female Labor Force Participation | 33.31% | 28.40% |
| % of Minorities with High Skill Job | 2.85% | 0.86% |

*Notes:* Innovation measured as the average number of patents per capita between 1880 and 1940. The top 10 states are: California, Connecticut, Delaware, Illinois, Massachusetts, Nevada, New Jersey, New York, Ohio, and Rhode Island. The bottom 10 states are: Alabama, Arkansas, Georgia, Mississippi, New Mexico, North Carolina, North Dakota, Oklahoma, South Carolina, and Tennessee.

Table 2: 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% |
| | | |
| Prob. Child: ≤ 35 yrs old | 72.9% | 80.0% |
| Prob. Child: > 35 yrs old | 80.9% | 89.7% |
| Av. # Children: ≤ 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 Born in Great Britain | 5.19% | 3.46% |
| Percent Born in Germany | 4.0% | 2.67% |
| Percent Born in Other Europe | 8.72% | 8.27% |
| Percent Born in Canada | 2.56% | 1.73% |
| Percent Born in Other Countries | 0.65% | 1.24% |
| | | |
| Percent Own Radio | 81.7% | 41.6% |
| Percent Of Population | 0.02% | 99.98% |

Figure 4: INVENTORS PER 10,000 BY RACE



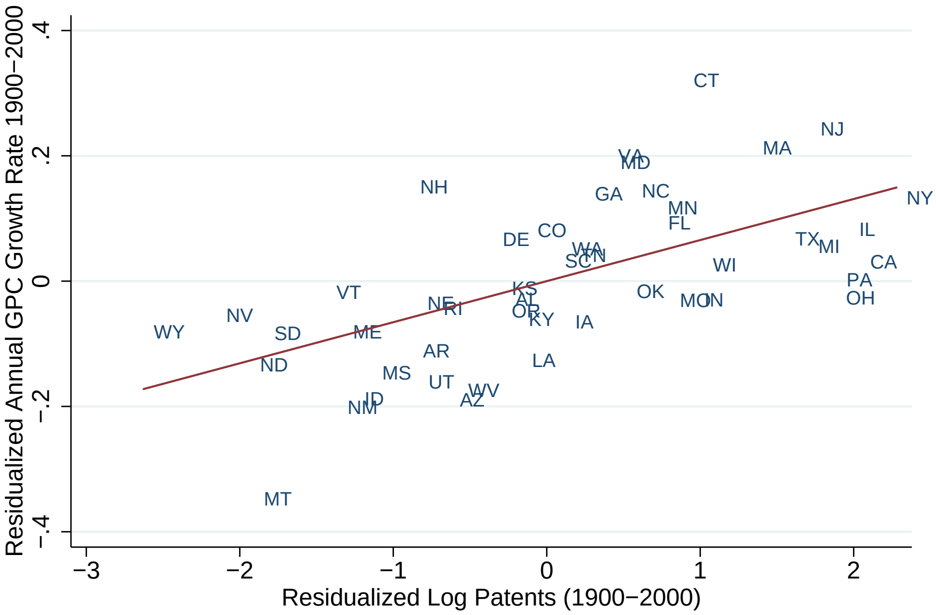
# 3 Empirical Analysis

## 3.1 Innovation and Economic Growth

The long-standing endogenous growth literature builds on the premise that long-run growth is driven by innovation and technological progress. Although this idea is intuitive, providing empirical support for this premise has been challenging due to data limitations on historical innovations. In fact, to our knowledge no study has documented a causal empirical relationship between innovation and growth outcomes for the US over the long run.

Figure 5 shows the basic correlation between a proxy measure of innovation (patents) and economic growth. To account for scale effects we plot variables residualized against 1900 log GDP per capita. The relationship is strongly positive. The positive relationship between innovation and output growth persists at the sector-level, as shown in Figure 6.

Figure 5: Innovation and Long-run Growth: US States between 1900-2000



Growth = 0.000 + 0.066 * Patents
Slope coefficient statistically significant at 1% level

Moreover, we show that the state-level result illustrated in Figure 5 is econometrically robust and can be verified with an instrumental variables strategy. Table 3 reports coefficients from growth regressions controlling for the long-run effects of initial conditions and population density. We find consistently positive and statistically significant effects in columns 1 and 2. These results are robust in columns 4 and 5 to applying a measurement approach used by Davis et al. (1996) in the employment literature that corrects for any potential bias associated with transitory shocks to growth and mean reversion.

The economic magnitude of these estimates is especially informative. To see this, consider Massachusetts (MA) versus Wyoming (WY). As shown in Figure 5, Massachusetts has been 4 times more innovative than Wyoming during the 20th century. Assume MA and WY had the same initial GDP per capita in 1900 and identical population densities. Our estimated coefficients imply that the gap

Figure 6: INNOVATION AND LONG-RUN GROWTH: 3-DIGIT SECTORS BETWEEN 1948-1986

Sector Growth(Residualized) = −9.464 + 2.585*Log Innovation(Residualized)
Slope coefficient statistically significant at 5% level

Table 3: INVENTIONS AND GROWTH: REGRESSIONS OF 1947-1987 GROWTH RATE ON INNOVATION

| | **Innovation measure: Log Patents** | | | | | |
|---|---|---|---|---|---|---|
| | Annualized Growth Rate | | | DHS Growth Rate | | |
| | OLS | OLS | IV | OLS | OLS | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log Patents | $0.14^{***}$ | $0.11^{***}$ | $0.14^{***}$ | $0.04^{***}$ | $0.04^{***}$ | $0.04^{***}$ |
| | (0.04) | (0.04) | (0.05) | (0.01) | (0.01) | (0.02) |
| Initial Log GDP per Capita | $-1.68^{***}$ | $-1.78^{***}$ | $-1.84^{***}$ | $-0.51^{***}$ | $-0.54^{***}$ | $-0.56^{***}$ |
| | (0.23) | (0.23) | (0.25) | (0.07) | (0.07) | (0.08) |
| Population Density | | $1.40^{**}$ | $1.24^{**}$ | | $0.44^{**}$ | $0.38^{**}$ |
| | | (0.65) | (0.58) | | (0.20) | (0.18) |
| Observations | 48 | 48 | 48 | 48 | 48 | 48 |
| Mean Growth | 2.50 | 2.50 | 2.50 | 0.91 | 0.91 | 0.91 |
| Std. Dev. of Growth | 0.44 | 0.44 | 0.44 | 0.13 | 0.13 | 0.13 |

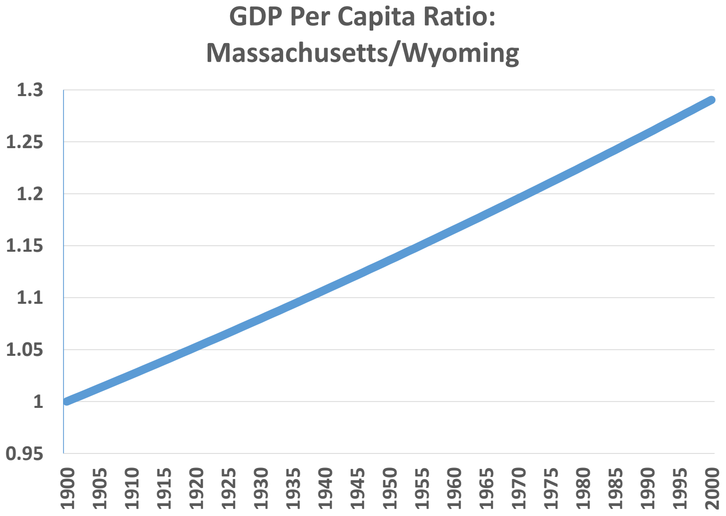
White heteroskedasticity robust standard errors reported in parentheses. DHS growth rate refers to the growth rate measure as proposed by Davis, Haltiwanger, and Schuh.

between MA and WY would have increased dramatically as illustrated in Figure 7. By the end of the century, MA would be 30% richer than WY just because of the differences in their innovativeness.

### 3.1.1 Instrumental Variables

In columns 3 and 6 of Table 3 we attempt to identify a causal effect in our growth regressions using OSRD contracts for wartime technological development as an instrument. Our strategy requires that these contracts were correlated with innovation, uncorrelated with omitted determinants, and

Figure 7: Economic Interpretation of Regression Results



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. A brief survey of the institutional setting for the OSRD along with quantitative placebo tests lends support to the credibility of our instrumentation approach.

The OSRD was established under President Roosevelt's Executive Order in June 1941 and operated extensively until it was terminated in December 1947. It was 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 fuze, navigation systems, solid fuel rockets, detonators and most famously the basic science used in the Manhattan Project (the Manhattan project was later transferred to the Manhattan District of the Army Engineers). Because of its significant impact, the OSRD spurred federal involvement in the development of American 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 specific 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 uncertainty, the OSRD contracted with multiple entities to solve the same problem simultaneously. 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, 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).[4]

We collected data on all contracts granted by the OSRD during its operation. We observe 1,717 contracts across 39 US 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 New York State 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 is 4.3 and the median is 1. The most prolific private firm in terms of contracting is the Western Electric Company with 107 contracts. The most prolific university is MIT which was granted 89 contracts. It received almost $117 million from the government during the war (Lowen, 1997, p.52).

Using these data, columns 3 and 6 of Table 3 reveal that are the OLS coefficients are confirmed by the corresponding IV estimates. To evaluate the validity of the exclusion restriction, we provide placebo tests of the instrument in Table 4. First, we check if our results change when we use 1935-1940 patents to predict the GDP growth rate. Our estimates in columns 1 and 2 are not statistically significant, indicating that the boost to growth we observe from innovation is not being driven by patenting activity just prior to the operation of the OSRD. Second, in columns 3 and 4 we check if contract allocation is correlated with pre-trend growth. Again, we do not find a statistically significant effect.

Table 4: PLACEBO TESTS

| **Dependent Variable:** | 1947-87 GDP Growth Rate | | Contracts | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Log Patents 1935-1940 | 0.115 | 0.006 | | |
| | (0.189) | (0.009) | | |
| 1935-1940 GDP Growth | | | 0.098 | |
| | | | (0.161) | |
| 1935-1940 GDP DHS Growth | | | | 2.107 |
| | | | | (3.386) |
| 1940 GDP per Capita | 0.389 | 0.018 | 3.129*** | 3.129*** |
| | (0.802) | (0.038) | (1.002) | (1.002) |
| Population Density | -6.731** | -0.318** | 12.582** | 12.587** |
| | (3.248) | (0.154) | (5.438) | (5.438) |
| Growth Rate | Annual | DHS | Annual | DHS |
| Observations | 48 | 48 | 48 | 48 |

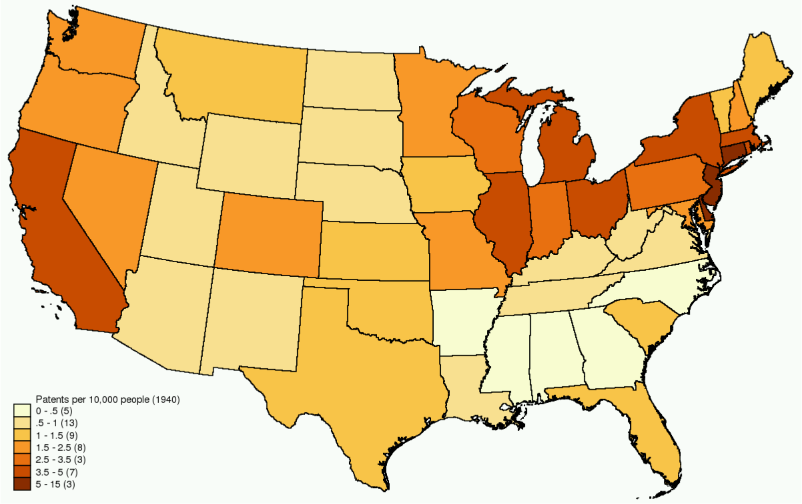
---

[4] Procurement related contracts generally “allowed for work on a fixed price plus a reasonable profit for the contractor” (Payne, 1992). In terms of rights to patents, the contractor generally retained these, though the government was permitted to a royalty-free compulsory license from the contractor or to a patent buyout at a reasonable price. If the contractor chose not to patent, the government retained the right to do so, and would in turn grant the contractor a non-exclusive royalty free license to use the invention (Wellerstein, 2008).

## 3.2 Regional Characteristics

Having established the empirical importance of innovation for long run growth, we now turn to understanding the determinants of differences in innovation outcomes. Following the road map outlined in Figure 2 we proceed to highlight important descriptive aspects of the innovation process.

Figure 8: THE GEOGRAPHY OF INVENTIVENESS: PATENTS PER 10,000 PEOPLE

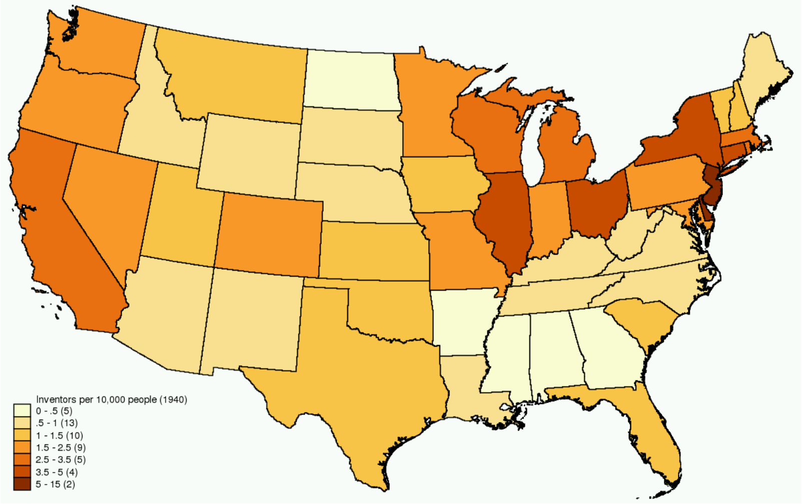


*Notes:* Darker colors represent more inventive regions.

Figures 8 and 9 illustrate the geography of inventiveness defined as patents and inventors per 10,000 people. 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 and this holds 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 US in 1900, it was ranked number 10 in 1920 and number 5 in 1940. Wyoming was innovative in several of our snapshot years, perhaps because of developments related to the evolution of the Union Pacific Railroad. However, as noted above, Wyoming under-performed relative to other states in GDP per capita terms over the very long run.

**Population Density and Urbanization** Table 1 illustrates the socio-economic profile of the top 10 and the bottom 10 inventive states. Population density is much higher in the most inventive states. As illustrated further in Figure 10, population density is positively correlated with patents per 10,000 people between 1880 and 1940. This finding is inline with two parallel literatures: First, a growing theoretical literature has argued that human interaction is key for human capital accumulation and economic growth (e.g., Lucas (2009)). Second, the agglomeration literature has long argued that physical proximity promotes creativity, the exchange of ideas and spillovers of knowledge capital among inventors (e.g., Glaeser (2011)). We also find that the top inventive states were associated with higher levels of education and higher skilled occupations. These results are consis-

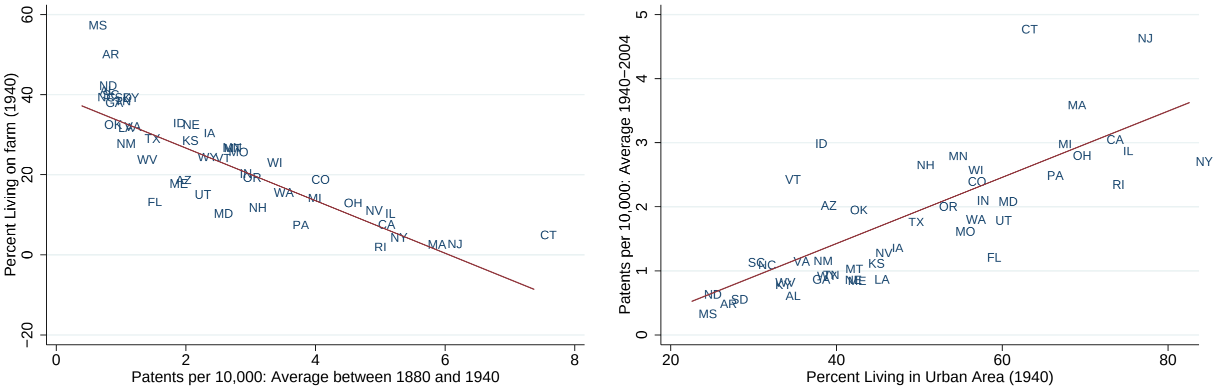
Figure 9: THE GEOGRAPHY OF INVENTIVENESS: INVENTORS PER 10,000 PEOPLE



*Notes:* Darker colors represent more inventive regions.

tent with denser places being more likely to create the type of externalities that lead to sustained economic growth.

Figure 10: POPULATION DENSITY



Urbanicity = 5.029 −0.109 * Patents per Capita
Slope coefficient statistically significant at 1% level

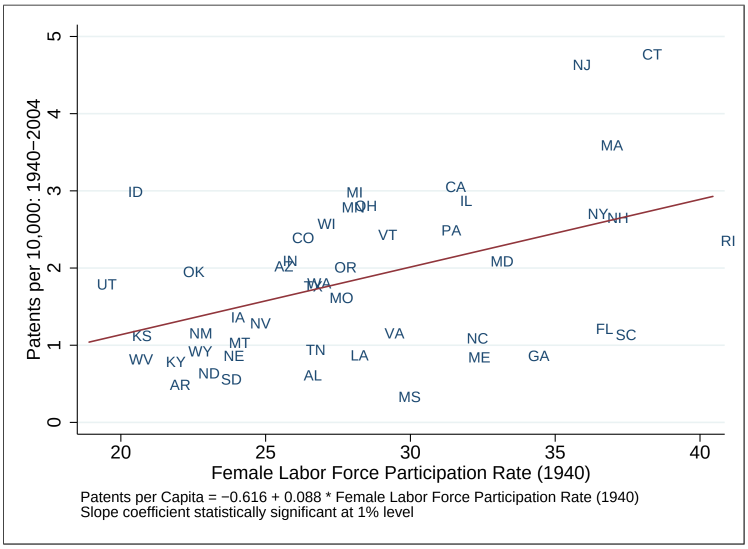
Patents per Capita = −0.640 + 0.052 * Percent Living in Urban Area
Slope coefficient statistically significant at 1% level

**Female Labor Force Participation** One potential explanation for these underlying state-level differences is that innovative places are relatively more open to unconventional and disruptive ideas. Insofar as this can be traced, in turn, to the makeup of society and the incentives for creative invention (e.g., Florida (2002), Acemoglu et al. (2014)), we would expect to find differences in the innovativeness of US states conditional on demographic characteristics.

One important dimension of being open to the changing world was the participation of women in the labor force. Goldin (2006) shows that women became increasingly active in the labor market during the early twentieth century as a consequence of innovations like the typewriter, which increased demand for clerical work, and a growth in female high school enrollment and graduation rates, which expanded workplace opportunities. At the same time there were still significant constraints. For example, “marriage bars” meant that by 1930, 60% of America’s public schools imposed some kind of restrictions on the employment of married women, while more than half fired women when they married. A gender-based division of labor persisted in firms, and it was not until the early 1940s that marriage bars were formally revoked.

To the extent that higher female labor force participation captures elements of a society’s openness to disruption, we should expect a positive association with the innovativeness of a region. Figure 11 shows supportive evidence in this direction when we correlate the female labor force participation rate in 1940 with long run technological development.

Figure 11: FEMALE LABOR FORCE PARTICIPATION



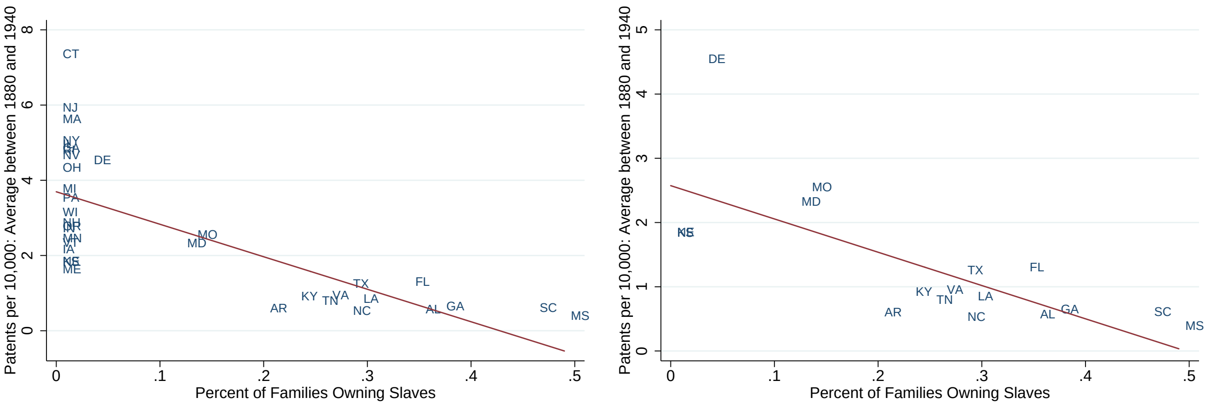
Patents per Capita = −0.616 + 0.088 * Female Labor Force Participation Rate (1940)
Slope coefficient statistically significant at 1% level

**Slave Ownership** Another interesting aspect of openness of a society could be seen from its approach toward slavery. Wright (1986) argues that because of its association with slavery, the southern economy of the US was constrained by a lack of technological innovation in agriculture and manufacturing. Slavery can undermine trust, having a persistent effect on beliefs and norms (Nunn and Wantchekon, 2011). A lack of cultural freedom to deviate from established norms can strongly inhibit innovation. 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.

We use data from the 1860 Census to determine the percentage of slave-owning families. To the extent that a low slave ownership rate capture the society’s openness to change, one would expect that innovation would correlate negatively with the slave ownership rate. This is indeed what we see in Figure 12, both when we correlate slave ownership with patenting in all US states and when we do this for only states in the southern part of the country, where slavery was most prevalent.

**Financial Development** Clearly one of the most important aspects of innovation is access to capital. There is a vast literature on this topic. Cross-country growth regressions have shown that higher levels of financial development are associated with faster rates of economic growth (e.g., King and Levine (1993), Rajan and Zingales (1998)). Within the US a range of evidence indicates

Figure 12: SLAVE OWNERSHIP



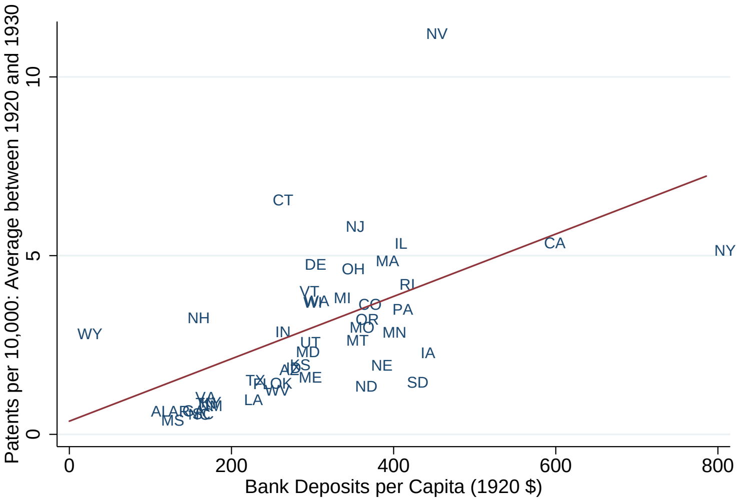
Patents = 3.695 −8.639 * Slave Ownership
Slope coefficient statistically significant at 1% level

Patents = 2.574 −5.183 * Slave Ownership
Slope coefficient statistically significant at 1% level

capital availability mattered for innovation. For example, Lamoreaux et al. (2004) find that venture-style intermediation of capital dramatically reduced financing constraints for inventors in Cleveland, an important second industrial revolution city.

Measuring the level of financial development is fraught with difficulties. Private transactions between investors and inventors are not observable systematically and most later stage R&D is financed by firms internally. However, we can gain useful insights using FDIC data, which provides broad indicators of financial market development. We use data on bank deposits since these reflect a measure of intermediation as savings are transformed into the extension of credit. Figure 13 shows a strong positive association between bank lending per capita and innovation.

Figure 13: DEPOSITS PER CAPITA (1920 $)

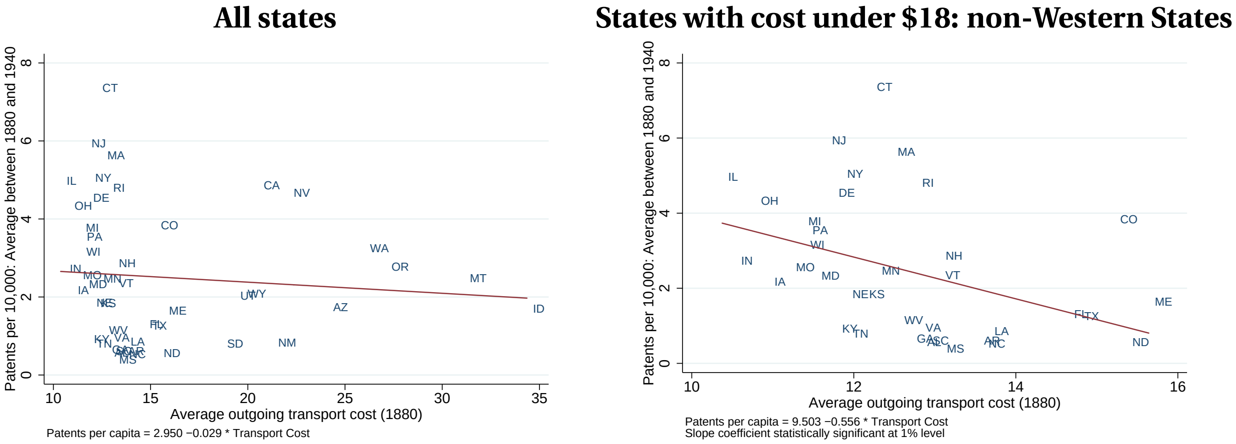


Patents = 0.370 + 0.009 * Deposits per Capita
Slope coefficient statistically significant at 1% level

**Geographical Connectedness and Railroads** Another important dimension for innovation is access to other geographical regions. This could both increase the market size for innovation and could also increase the potential flow of knowledge and spillovers. Both mechanisms receive support in the literature. Sokoloff (1988) found that inventive activity accelerated in locations that were proximate to navigable waterways, while Perlman (2016) finds strong effects on invention and agglomeration from the nineteenth century development of railroads. Donaldson and Hornbeck (2016) measure the increased level of market access caused by an expansion of the US railroad network between 1870 to 1890, finding the aggregate impact to be large.

Figure 14 investigates the relationship between average outgoing transportation costs from a region and its level of innovativeness. While the left panel focuses on all states, the right panel focuses on the ones below a certain cost threshold. It turns out that this nonlinearity is important. Among the states where the average outgoing transportation cost is below a certain threshold, the relationship between innovation and transportation cost is strongly negative, suggesting that access to outside markets was an important component of innovation.

Figure 14: TRANSPORT COST STATE SCATTERS



**All states**

**States with cost under $18: non-Western States**

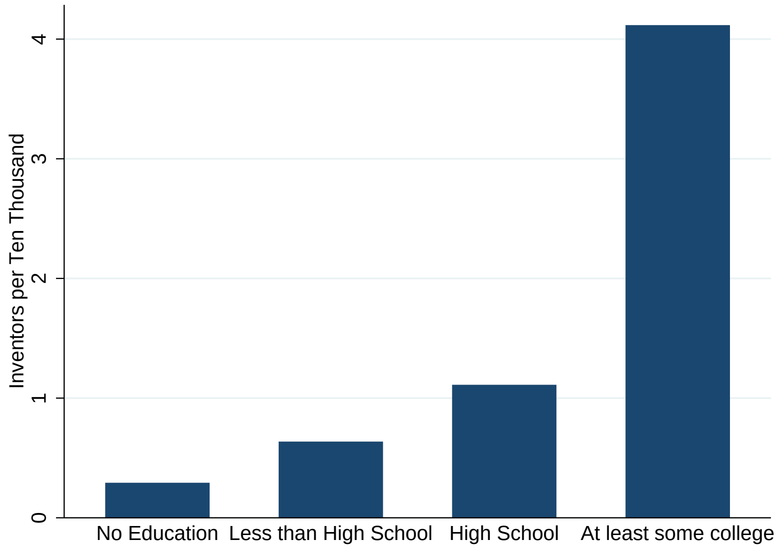
*Notes:* The horizontal axis plots the average cost to transport one ton of goods from a county in state $s$ to counties in state $d$ different from $s$. For details on the construction of the county-to-county transportation cost measures, see Donaldson and Hornbeck (2016).

## 3.3 Who Became an Inventor?

**Education and Parental Background** Next we turn to the role of education for inventiveness. There is a very large literature on the role of education for economic growth (see, for instance, Lucas (1998), Benhabib and Spiegel (1994), Bils and Klenow (2000), Goldin and Katz (2009), Barro (2001), Vandenbussche et al. (2006), Stokey (1991)). One of the main channels through which education affects economic growth may be its impact on innovation. Figure 15 shows the number of inventors per 10,000 people within each education group. While education seems to be an important determinant of becoming an inventor, the effect is particularly strong at the college degree level. For example, an individual with at least a college degree is 4 times more likely to become an inventor than an individual with just a high-school diploma.

Figure 16 illustrates the relationship between father's income and the probability of becoming an inventor. We find a strong association between the two, especially for the highest-income fathers. This suggests that if education was an important determinant of innovation as suggested by the

Figure 15: EDUCATION AND PROBABILITY OF BECOMING AN INVENTOR



previous figure, then the fact that only wealthy individuals had access to education could imply that credit constraints may have been binding (e.g., Celik (2015)). Since inventors need to raise capital to develop their ideas, this would be one channel influencing entry into inventive activity.

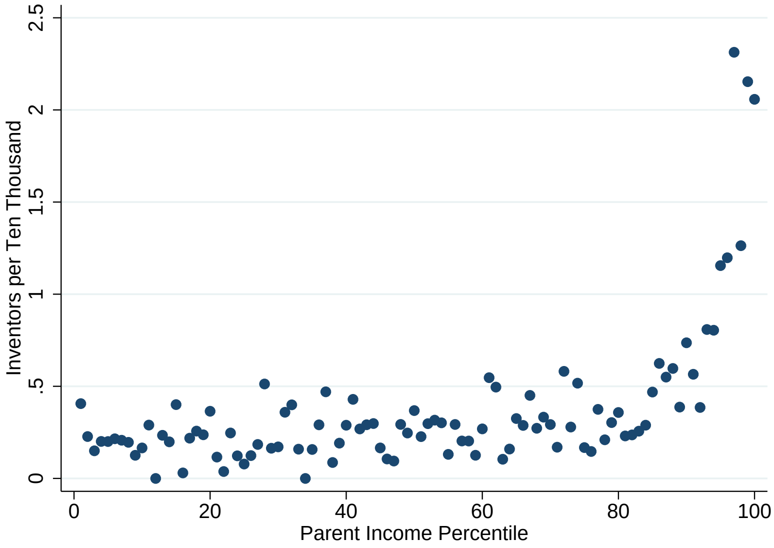
Table 5 examines this set of relationships more systematically using linear probability models. The dependent variable in these regressions is an indicator for being granted at least one patent, scaled by a factor of 100 for legibility. Column 1 establishes a positive correlation between the father being an inventor and the child being an inventor. Column 2 introduces parental income instead and column 3 includes both measures. A father's income is still very strongly correlated with the child becoming an inventor.

Of course, a potentially confounding effect is that high-income parents could be highly-educated parents who invest more in their children's development. To address this, column 4 adds parental education. Interestingly, parental income still matters. Finally, in column 5 we include the child's own education. The effect of parental income disappears, which suggests that parental income only positively affects the probability of becoming an inventor through its effect on childrens' access to education.

While Table 5 focuses on the extensive margin—the characteristics of those becoming inventors—Table 6 extends the analysis to consider the effect of background on the number of career patents and citations generated amongst the universe of inventors (i.e., the intensive margin). In column 1 we find a weak positive effect of the father being an inventor, which is also statistically significant in the citations specification. In columns 2, 3 and 4 we do not detect a strong effect of father's income, or father's education. In column 5 we introduce the child's own education and this is strongly correlated with long run inventiveness. In other words, the most highly educated inventors tended to be the most productive.

Two findings emerge when taking Table 5 and Table 6 together. First, the importance of education holds both at the extensive and intensive margins, which is consistent with a human capital

Figure 16: THE RELATIONSHIP BETWEEN FATHER'S INCOME AND THE PROBABILITY OF BECOMING AN INVENTOR



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 business and management practices literatures, showing that privileged access to career paths (e.g., inherited CEO roles in family firms) is associated with under performance (e.g., Perez-Gonzalez (2006); Bloom and Van Reenen (2007) Caselli and Gennaioli (2013)).

**Timing of Innovation** We find that inventors are most productive between ages 35-55 as illustrated in Figure 17. In that sense they had a reasonably long productive career-length. Long career-length is consistent with Khan and Sokoloff (2004)'s data on superstar inventors. They found that while 37 percent born prior to 1820 had careers over 30 years, 57 percent did in their post-1820 birth cohorts. The age profile of innovation matters for economic growth because a broad inventor life cycle, like we observe, tends to maximize creative output (Jones (2010)).

**Family Choices** This life cycle dynamic created tradeoffs. While we have already established that education (costly in terms of time) was an important input into the process of becoming an inventor, Figure 18 shows that inventors delayed marriage substantially and had fewer children on average. In some sense these two findings are mechanical; as Becker (1974, p.22) pointed out, "the age of entry [into marriage] would be earlier the larger the number of children desired." Yet, the type of behavior we observe can also be explained by theoretical models of marriage markets like Bergstrom and Bagnoli (1993), in which 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.

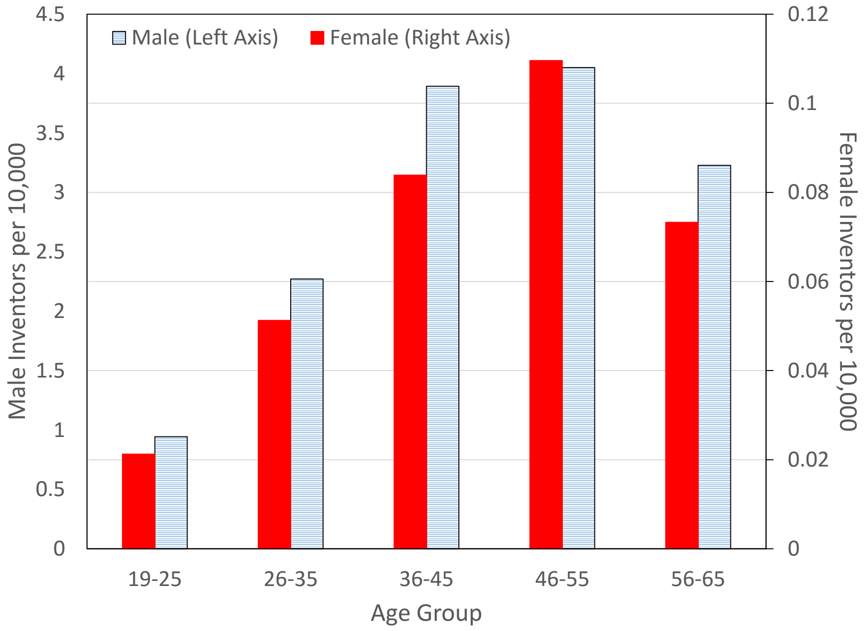
Anecdotally, some of the most prolific inventors did not believe in marriage at all. Nikola Tesla thought that marriage was inconsistent with great invention. He famously commented in the *New York Herald* in 1897 "I do not believe an inventor should marry, because he has so intense a nature,

Table 5: Who Becomes and Inventor? Regressions on the Inventor indicator

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Father Inventor | $0.161^{**}$ | | $0.159^{**}$ | $0.157^{**}$ | $0.155^{**}$ |
| | (0.075) | | (0.076) | (0.075) | (0.075) |
| Father Income $90^{th}-95^{th}$ %ile | | $0.003^{**}$ | $0.003^{**}$ | $0.002^{*}$ | -0.000 |
| | | (0.001) | (0.001) | (0.001) | (0.001) |
| Father Income $95^{th}$ %ile and above | | $0.008^{***}$ | $0.008^{***}$ | $0.006^{***}$ | 0.001 |
| | | (0.002) | (0.002) | (0.002) | (0.002) |
| Father: High School Graduate | | | | $0.004^{**}$ | -0.001 |
| | | | | (0.001) | (0.001) |
| Father: At least Some College | | | | $0.007^{***}$ | $-0.002^{*}$ |
| | | | | (0.001) | (0.001) |
| Self: High School Graduate | | | | | $0.006^{***}$ |
| | | | | | (0.001) |
| Self: At least Some College | | | | | $0.029^{***}$ |
| | | | | | (0.004) |
| Observations | 82810280 | 82810280 | 82810280 | 82810280 | 82810280 |
| R-squared | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 |

Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, and a quadratic in 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 category is below median income.

Figure 17: Probability of Innovation by Age



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 in fact marry. For example, Elias Howe, who invented the sewing machine, married when he was 21 years of age. Thomas Edison married first at age 24 and within a

Table 6: INDIVIDUAL BACKGROUND AND LONG RUN INVENTIVENESS: REGRESSIONS ON THE INTENSIVE MARGIN

| | **Panel A: Log Career Patents** | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Father Inventor | 0.19 | | 0.21 | 0.16 | 0.02 |
| | (0.74) | | (0.70) | (0.63) | (0.63) |
| Father Income $90^{th} - 95^{th}$ %ile | | -0.27 | -0.27 | -0.23 | -0.24 |
| | | (0.23) | (0.23) | (0.24) | (0.24) |
| Father Income $95^{th}$ %ile and above | | 0.14 | 0.14 | 0.03 | 0.01 |
| | | (0.20) | (0.20) | (0.20) | (0.19) |
| Father: High School Graduate | | | | 0.07 | -0.03 |
| | | | | (0.11) | (0.12) |
| Father: At least Some College | | | | 0.23 | 0.12 |
| | | | | (0.14) | (0.13) |
| Self: High School Graduate | | | | | 0.06 |
| | | | | | (0.04) |
| Self: At least Some College | | | | | $0.30^{***}$ |
| | | | | | (0.05) |
| Observations | 9032 | 9032 | 9032 | 9032 | 9032 |
| Mean of Dep. Var. | 1.58 | 1.58 | 1.58 | 1.58 | 1.58 |
| S.D. of Dep. Var. | 1.37 | 1.37 | 1.37 | 1.37 | 1.37 |

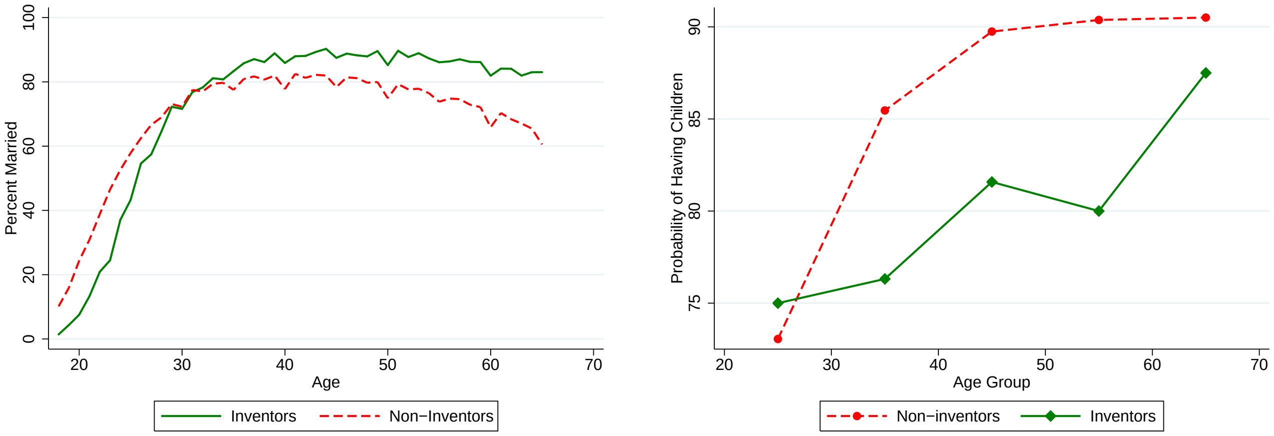
| | **Panel B: Log Career Citations** | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Father Inventor | $1.19^{*}$ | | $1.19^{*}$ | $1.07^{*}$ | 0.89 |
| | (0.67) | | (0.63) | (0.56) | (0.59) |
| Father Income $90^{th} - 95^{th}$ %ile | | -0.23 | -0.24 | -0.20 | -0.21 |
| | | (0.26) | (0.27) | (0.30) | (0.30) |
| Father Income $95^{th}$ %ile and above | | 0.27 | 0.27 | 0.19 | 0.18 |
| | | (0.26) | (0.25) | (0.20) | (0.19) |
| Father: High School Graduate | | | | $-0.38^{*}$ | $-0.51^{**}$ |
| | | | | (0.21) | (0.22) |
| Father: At least Some College | | | | 0.21 | 0.07 |
| | | | | (0.22) | (0.23) |
| Self: High School Graduate | | | | | 0.03 |
| | | | | | (0.06) |
| Self: At least Some College | | | | | $0.37^{***}$ |
| | | | | | (0.06) |
| Observations | 9032 | 9032 | 9032 | 9032 | 9032 |
| Mean of Dep. Var. | 2.82 | 2.82 | 2.82 | 2.82 | 2.82 |
| S.D. of Dep. Var. | 1.84 | 1.84 | 1.84 | 1.84 | 1.84 |

Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, and a quadratic in 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 category is below median income.

year had developed the revolutionary quadruplex telegraph for sending messages simultaneously over a single wire. Following the death of his first wife Edison married again at age 39. Figure 18 shows that inventors did indeed marry at a higher rate than their non-inventor counterparts, at least at older ages.

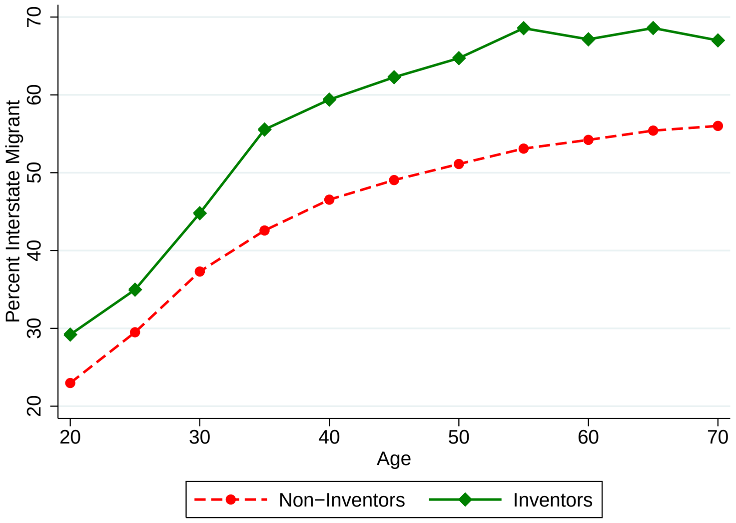
**Migration** Inventors were not a static group. We find higher levels of interstate and international migration in the inventive states. In particular we find a much higher share of international mi-

Figure 18: PROBABILITY OF BEING MARRIED AND HAVING KIDS



grants in the top 10 inventive states which is in line with Akcigit et al. (2016b)'s finding that inventors are internationally highly mobile. Although modern studies have produced opposing conclusions on the role of immigrants in determining levels of inventiveness (e.g., Kerr and Lincoln (2010); Hunt and Gauthier-Loiselle (2010)), historical evidence is more unequivocal. For example, Moser et al. (2014) estimate that German emigres who fled the Nazi regime provided a significant boost to US invention during the twentieth century. Figure 19 shows the migration rates by age. What we see is that inventors were most likely to move after the age of 35 which is the the beginning of their most innovative period according to Figure 17.

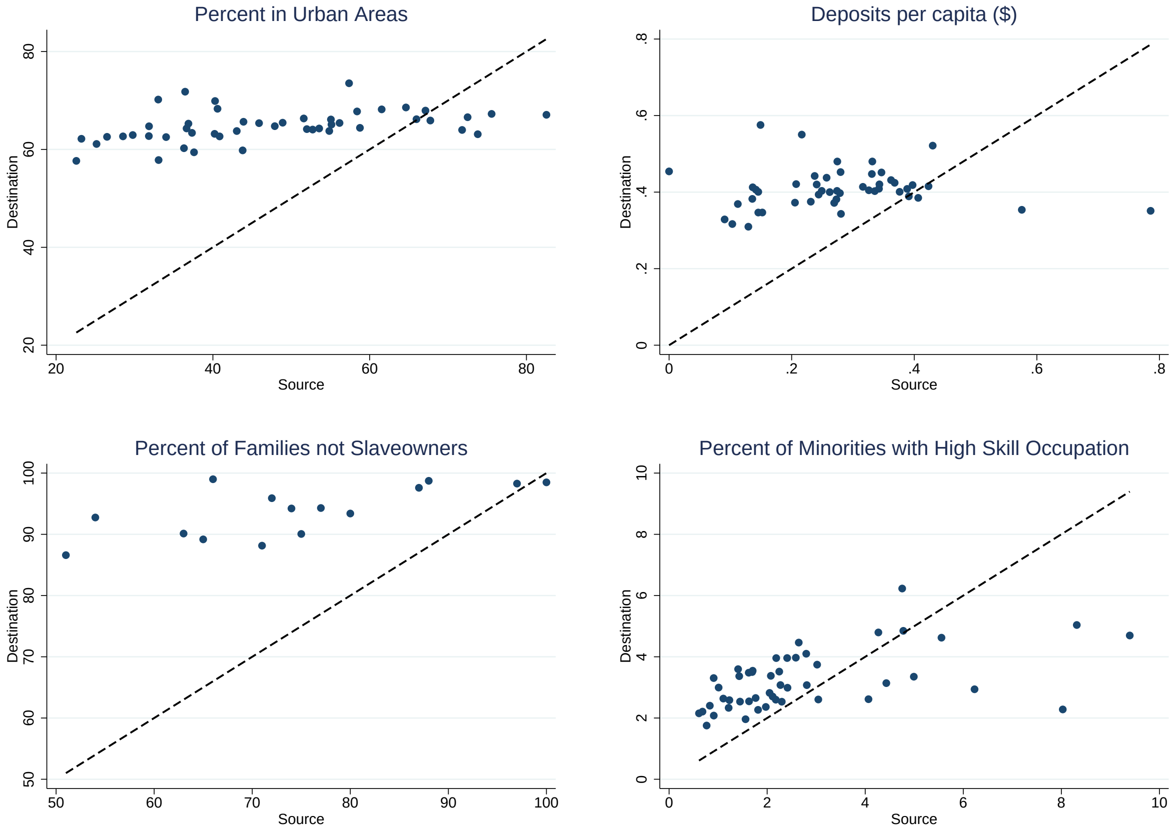
Figure 19: INTERSTATE MIGRATION RATES BY AGE



Conditional on moving to a new location, where did inventors go? To answer this question, Figure 20 below plots the characteristics of geographical origin and destination amongst inventors who move across state lines. To facilitate the exposition, a 45-degree line is also plotted which denotes no change at all. The top-left panel shows that most of the observations are clustered above the

45-degree line implying that the inventors were moving from less to more urbanized regions. Likewise, the top-right panel shows that inventors were moving toward regions where deposit ratios were higher, suggesting that access to finance could have played a role in their migration decisions. The bottom-left and bottom-right panels show that inventors moved toward regions where slave-ownership had been lower and the fraction of minorities with high-skill occupations was higher. We argue that this evidence is consistent with inventors moving towards regions that were more open to disruption.

Figure 20: WHERE DO THE INVENTORS MOVE?



## 3.4 Innovation and Labor Income

Some of the most valuable information in the 1940 census relates to labor income. Although the information is not recorded for all observations we are able to construct data on labor income sufficiently well to allow us to study the relationship between innovation and income. First, our data shows a very strong correlation between the citation-weighted patent portfolio of an inventor and log wages. This is important because it is often argued that inventors have intrinsic preferences to the point where inventors have a “taste” for technological development rather than financial payoffs (e.g., Stern (2004)). Our evidence suggests that the two were not incompatible and may have reflected efficiency. Occupational choice models (e.g., Roy (1951)) predict that individuals will sort into occupations that provide them with the highest expected earnings.

Second, Figure 22 reveals that inventors had higher wages than non-inventors and the residuals from a Mincer regression are also to the right of non-inventors. Since we have a rich vector of co-variates in these regressions and the residuals reflect unmeasured attributes, these figures provide suggestive evidence that invention was a key labor income differentiator.

Figure 21: THE RELATIONSHIP BETWEEN CITATIONS AND WAGES

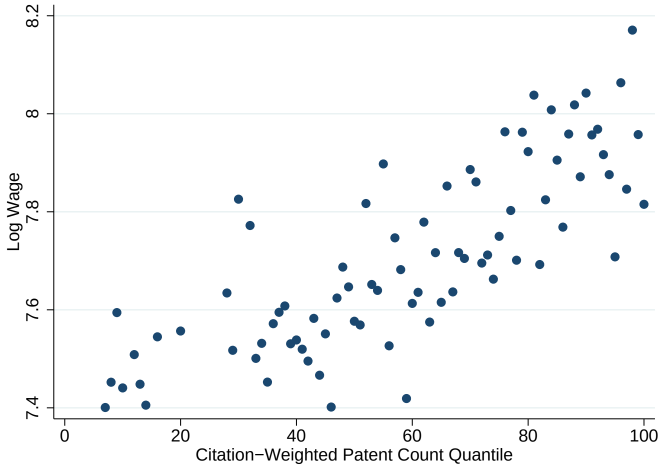
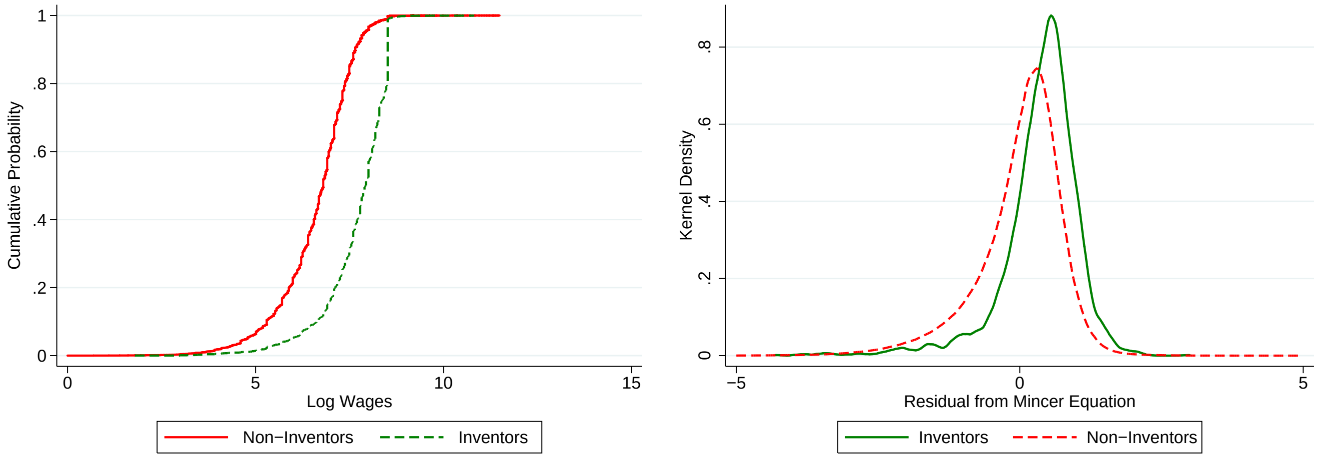


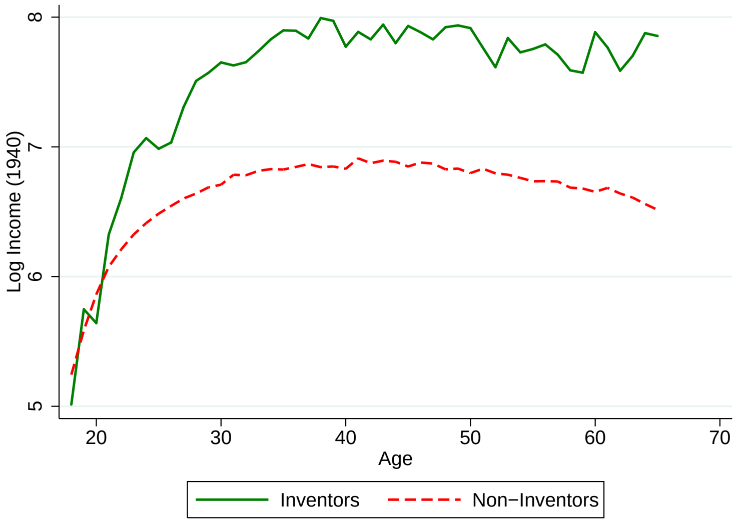
Figure 22: THE DISTRIBUTION OF LABOR INCOME BY INVENTOR STATUS (1940)



## 3.5 Income Inequality and Social Mobility

**Income Inequality** We also exploit labor income data in the 1940 census to study regional income dynamics. This issue is especially relevant given the large literature on income inequality and recent attempts to analyze the relationship between the top income share and patenting. Aghion et al. (2015a) examine modern US 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

Figure 23: THE LIFE CYCLE OF EARNINGS BY INVENTOR STATUS



relationship between inequality and patenting becomes much weaker at different thresholds like the top 10% share, and they even find a negative relationship when using the Gini coefficient, which considers all parts of the income distribution not just the top share.

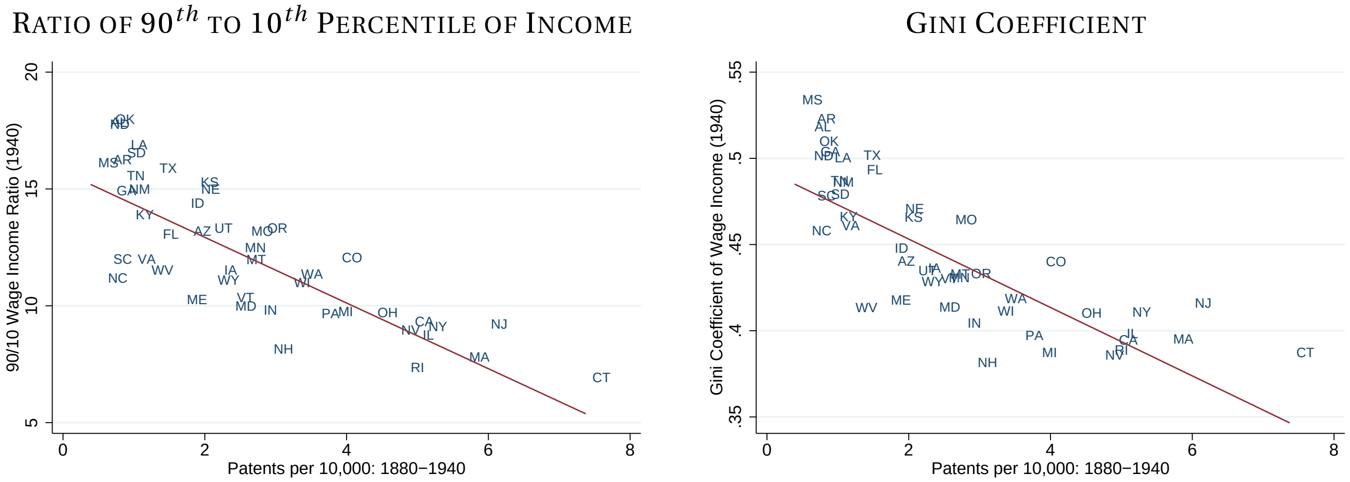
Our results in Figure 24 generally point to a negative relationship between income inequality and inventiveness.[5] We calculate the 90/10 ratio and Gini coefficient, both of which are very strongly negatively associated with regional innovativeness. However, we also find some sensitivity to measurement. The top 1% income share exhibits a non-linear U-shaped relationship with patenting. In the least innovative states we find a negative relationship. However, in the most innovative states like New York, New Jersey and Massachusetts we find that more patenting was associated with more income held in the top 1% share.

**Social Mobility** One explanation for this result is that innovation could be a significant social elevator. Indeed, we see that more innovative regions feature more social mobility, defined in Figure 25 as the fraction of high-skill people conditional on having a low-skill father. We find in Table 7 that the relationship between patents per capita and social mobility is positive and statistically significant. These results are especially important from the standpoint of innovation as a driver of social change given that Long and Ferrie (2013) find that around the turn of the twentieth century America was (from its mid-nineteenth century high-point) generally becoming less socially mobile.

---

[5] See Aghion et al. (2015a) and Jones and Kim (2014) for theoretical treatments of inequality and innovation.

Figure 24: RELATIONSHIP BETWEEN WAGE INCOME INEQUALITY AND INVENTIVENESS



90/10 Wage Income Ratio (1940) = 15.742 −1.405 * Patents per Capita
Slope coefficient statistically significant at 1% level

Gini Coefficient of Wage Income (1940) = 0.493 −0.020 * Patents per Capita
Slope coefficient statistically significant at 1% level

SHARE OF INCOME HELD BY TOP 1%

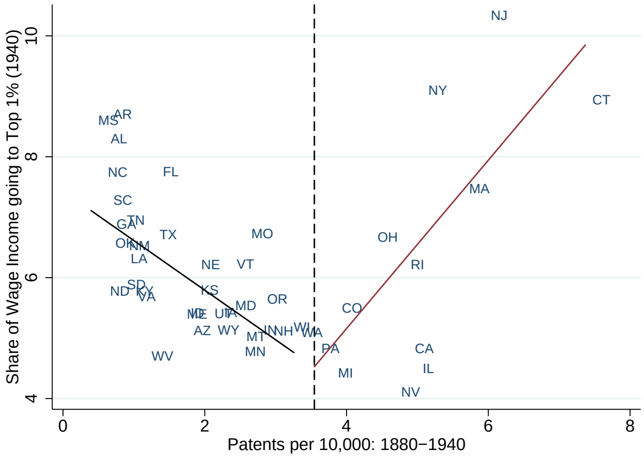
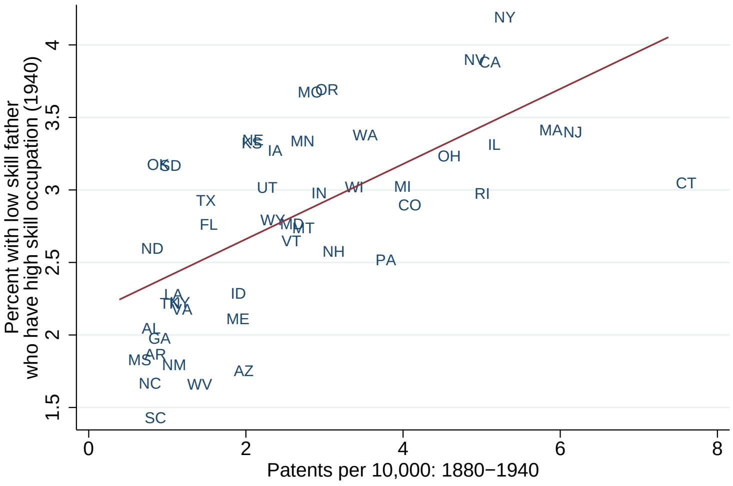


Figure 25: THE RELATIONSHIP BETWEEN INVENTIVENESS AND SOCIAL MOBILITY



Social Mobility = 2.143 + 0.259 * Patents per Capita
Slope coefficient statistically significant at 1% level

Table 7: % OF HIGH-SKILL CHILD GIVEN LOW-SKILL FATHER

| | (1) | (2) |
|---|---|---|
| Av. Patents per Capita 1880-1940 | $0.380^{***}$ | $0.292^{***}$ |
| | (0.048) | (0.089) |
| Confederate State | | -0.122 |
| | | (0.153) |
| % Agricultural Occupation (1940) | | -0.015 |
| | | (0.011) |
| % Manufacturing Occupation (1940) | | -0.010 |
| | | (0.017) |
| Observations | 48 | 48 |
| R-squared | 0.7178 | 0.7334 |
| Mean of Dep. Var. | 2.13 | 2.13 |
| Std. Dev. of Dep. Var. | 0.78 | 0.78 |

## 4 Conclusion

This paper presents new facts emerging from a major data collection and matching exercise combining information from US patent records with data on individuals from Federal Censuses between 1880 and 1940. The new data provide a comprehensive profile of inventions and their creators during the golden age of US invention. Our analysis began with an attempt to identify a causal relationship between innovation and long run economic growth. We proceeded to explore some of the main mechanisms driving this relationship using a framework for establishing macro and micro-level facts about the innovation process. While we present numerous interesting correlations in this paper, establishing causal links between our variables of interest is left to future research.

We believe that examining the drivers of innovation during this historical time period is critical because these data have the potential to shed light on numerous key debates on innovation and long-run economic growth. They can also complement modern studies such as Aghion et al. (2015b) and Bell et al. (2015) to provide a more complete picture of comparisons and contrasts over time. So far, we have focused on regional performance, openness to disruption, financial development, the returns to innovation and the relationship between social mobility and invention. These areas have allowed us to gain preliminary insights into the birth of technological ingenuity during one of the most important eras of American economic development.

## References

**Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson**, “A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration,” *Journal of Political Economy*, 2014, *122* (3), 467–506.

**Acemoglu, Daron, Ufuk Akcigit, and Murat Alp Celik**, "Young, Restless and Creative: Openness to Disruption and Creative Innovations," 2014. National Bureau of Economic Research WP # 19894.

**Aghion, Philippe and Peter Howitt**, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2).

**_ , Ufuk Akcigit, and Peter Howitt**, "What Do We Learn from Schumpeterian Growth Theory?," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, 2014, pp. 515–563.

**_ , _ , Antonin Bergeaud, Richard Blundell, and David Hémous**, "Innovation and Top Income Inequality," 2015. NBER Working Paper # 21247.

**_ , _ , Ari Hyytinen, and Otto Toivanen**, "Living the American Dream in Finland: The Social Mobility of Innovators," 2015. University of Chicago mimeo.

**Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood**, "Buy, Keep or Sell: Economic Growth and the Market for Ideas," *Econometrica*, 2016, *84* (3), 943–984.

**_ , Salome Baslandze, and Stefanie Stantcheva**, "Taxation and the International Migration of Inventors," *American Economic Review*, 2016. forthcoming.

**Barro, Robert J**, "Human Capital and Growth," *American Economic Review*, 2001, *91* (2), 12–17.

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

**Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen**, "The Lifecycle of Inventors," 2015. Harvard mimeo.

**Benhabib, Jess and Mark M Spiegel**, "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-country Data," *Journal of Monetary Economics*, 1994, *34* (2), 143–173.

**Bergstrom, Theodore C and Mark Bagnoli**, "Courtship as a Waiting Game," *Journal of Political Economy*, February 1993, *101* (1), 185–202.

**Bils, Mark and Peter J Klenow**, "Does Schooling Cause Growth?," *American Economic Review*, 2000, pp. 1160–1183.

**Bloom, Nicholas and John Van Reenen**, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

**Caselli, Francesco and Nicola Gennaioli**, "Dynastic Management," *Economic Inquiry*, 2013, *51* (1), 971–996.

**Celik, Murat Alp**, "Does the Cream Always Rise to the Top? The Misallocation of Talent and Innovation," 2015. University of Pennsylvania Mimeo.

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

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

**Davis, Steven, John Haltiwanger, and Scott Schuh**, "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts," 1996.

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

**Feigenbaum, James J**, "A New Old Measure of Intergenerational Mobility: Iowa 1915 to 1940," 2015. Harvard mimeo.

**Field, Alexander J**, "The Most Technologically Progressive Decade of the Century," *American Economic Review*, 2003, *93* (4), 1399–1413.

**Florida, Richard**, "The Economic Geography of Talent," *Annals of the Association of American geographers*, 2002, *92* (4), 743–755.

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

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

**Goldin, Claudia Dale and Lawrence F Katz**, *The Race between Education and Technology*, Harvard University Press, 2009.

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

**Hunt, Jennifer and Marjolaine Gauthier-Loiselle**, "How Much Does Immigration Boost Innovation?," *American Economic Journal: Macroeconomics*, 2010, pp. 31–56.

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

**Jones, Charles I and Jihee Kim**, "A Schumpeterian Model of Top Income Inequality," 2014. NBER Working Paper # 20637.

**Kerr, William R and William F Lincoln**, "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention," *Journal of Labor Economics*, 2010, *28* (3), 473–508.

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

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

**Klein, Alexander**, "New State-Level Estimates of Personal Income in the United States, 1880-1910," in Christopher Hanes and Susan Wolcott, eds., *Research in Economic History (Volume 29)*, Emerald Group Publishing Limited, 2013, pp. 191–255.

**Lamoreaux, Naomi R and Kenneth L Sokoloff**, "Inventive activity and the market for technology in the United States, 1840-1920," 1999. National Bureau of Economic Research WP # 7107.

**Lamoreaux, Naomi R., Margaret Levenstein, and Kenneth L. Sokoloff**, "Financing Invention During the Second Industrial Revolution: Cleveland, Ohio, 1870-1920," 2004. NBER Working Paper # 10923.

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

**Lowen, Rebecca**, *Creating the Cold War University: The Transformation of Stanford*, Berkeley: University of California Press, 1997.

**Lucas, Robert E**, "On the Mechanics of Economic Development," *Econometric Society Monographs*, 1998, *29*, 61–70.

_ , "Ideas and Growth," *Economica*, 2009, *76* (301), 1–19.

**Moser, Petra, Alessandra Voena, and Fabian Waldinger**, "German-Jewish Emigres and US Invention," *American Economic Review*, 2014, *104* (10), 3222–55.

**Mowery, David C.**, "Military R&D and Innovation," in "Handbook of the Economics of Innovation" 2010, chapter 29, pp. 1219–1256.

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

**Nunn, Nathan and Leonard Wantchekon**, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, December 2011, *101* (7), 3221–52.

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

**Perez-Gonzalez, Francisco**, "Inherited Control and Firm Performance," *American Economic Review*, December 2006, *96* (5), 1559–1588.

**Perlman, Elisabeth R.**, "Dense Enough To Be Brilliant: Patents, Urbanization, and Transportation in Nineteenth Century America," Working Paper, Boston University January 2016.

**Rajan, Raghuram and Luigi Zingales**, "Financial Dependence and Growth," *American Economic Review*, 1998, *88* (3), 559–586.

**Romer, Paul Michael**, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–102.

**Roy, A. D.**, “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 1951, *3* (2), 135–146.

**Sokoloff, Kenneth L.**, “Inventive Activity in Early Industrial America: Evidence From Patent Records, 1790–1846,” *Journal of Economic History*, December 1988, *48* (04), 813–850.

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

**Stern, Scott**, “Do Scientists Pay to Be Scientists?,” *Management Science*, June 2004, *50* (6), 835–853.

**Stokey, Nancy L**, “Human Capital, Product Quality, and Growth,” *Quarterly Journal of Economics*, 1991, *106* (2), 587–616.

**Vandenbussche, Jérôme, Philippe Aghion, and Costas Meghir**, “Growth, Distance to Frontier and Composition of Human Capital,” *Journal of Economic Growth*, 2006, *11* (2), 97–127.

**Wellerstein, Alex**, “Patenting the Bomb: Nuclear Weapons, Intellectual Property, and Technological Control,” *Isis*, 2008, *99* (1), 57–87.

**Wright, Gavin**, *Old South, New South: Revolutions in the Southern Economy since the Civil War*, New York: Basic Books, 1986.