Compulsory Innovation: How Education Forms Inventors

Colin Davison[1], Kirk Doran[2,3], and Chungeun Yoon[4]

[1]College of Wooster

[2]University of Notre Dame

[3]IZA - Institute of Labor Economics

[2]KDI School of Public Policy and Management

November 9, 2023

Click Here for Current Version

Abstract

We investigate the effect of education on innovation using the rollout of compulsory schooling laws in the United States. We find that primary and secondary education increase both the chance an individual will become an inventor and that they will patent a breakthrough innovation. The positive effect of formal schooling on the probability of invention is strongest in geographies with less pre-existing innovative activity, suggesting that exposure to innovation and education act as substitutes in the production of inventors. Our findings provide empirical evidence for the long-hypothesized link between education and the human capital necessary for innovation and growth.

1 Introduction

For the past several decades, one of the most important literatures in empirical economics has consistently supported the view that education yields long-run benefits for students and the economy. However, we still know little about the role of formal education in encouraging new ideas and innovation. This is surprising, given that the idea that education can lead to greater innovation is at the foundation of our most important models of economic growth, and is consonant with a long and varied literature on the effects of schooling on abilities and labor market outcomes. But, in spite of the a priori arguments in its favor, the actual evidence in favor of a causal relationship between formal education and innovation is scarce, and many of the most famous examples in history point in the opposite direction. For example, the Wright Brothers, inventors of the first airplane, did not graduate high school. Thomas Edison had little formal education in his childhood. Steve Jobs, Bill Gates, and Mark Zuckerberg dropped out of school and became successful inventors, entrepreneurs and innovators. Are these examples typical, or is there a causal relationship between an individual's formal education and their probability of invention?

We explore the question of how education impacts the formation of inventors by linking individuals in the Full Count 1940 U.S. Census with the patent record in order to track an individual's lifetime innovative output. We combine this data with information on compulsory schooling laws for the 1885-1912 birth cohorts (Clay et al. 2021). These cohorts experienced the rollout of laws designed to increase the educational attainment of children up to the age of 14. Using compulsory schooling laws as an instrument for years of schooling achieved, we are able to uncover exogenous variation in the amount of schooling individuals received based on their state and year of birth.

We start by examining the effects of compulsory schooling on the years of schooling an individual obtains. Using a Poisson model, we find that requiring seven or more years of schooling has a positive effect on the years of schooling an individual obtains. On average, when students are required to complete nine years of schooling this increases their educational attainment by 2%, relative to no compulsory schooling requirement. Estimating the same model but with lifetime patenting counts as the dependent variable we find that requiring nine years of schooling leads to a 10% increase in patenting. The results are robust to using breakthrough innovations as the dependent variable, suggesting that the increased educational attainment led to valuable innovation.

Next, we use compulsory schooling laws as an instrument for educational attainment in a two stage least squares (2SLS) estimation strategy, and we find that increased education raises the probability an individual becomes an inventor. We find that a marginal year of education increases the probability of becoming an inventor by 31-45% off the mean probability. The results are similar when we examine the effect of education on individuals becoming breakthrough inventors. When we allow years of schooling to remain endogenous, we arrive at smaller estimates, in which the marginal year of schooling is associated with a 13% increase in the probability of becoming an inventor off the mean level. Our results support the view that more years of formal education, even at the lower end of the schooling distribution, can increase the probability a person becomes an inventor and that they go on to produce significant quantities of useful innovative output.

Next, in regard to the formation of inventors, we test whether early childhood exposure to innovation acts as a complement or substitute to education. There are compelling reasons why either may be the case. In favor of a complementary view is the idea that early exposure to innovation is necessary for demonstrating the possibility of such an activity and inspiring children to consider it, but that education is still needed to attain the skills needed to take advantage of this exposure to innovation. On the other hand, individuals growing up in areas with little innovative activity may have the most to gain from increased education, since in these locations there is less of a possibility of picking up on tacit knowledge through their activities outside of school. To distinguish between these stories, we examine if the effect of education on the formation of inventors is stronger or weaker for individuals born in states with more patenting per capita, manufacturing output per capita, or urban share. We find that the effectiveness of education in forming inventors declines with innovative and economic activity. This result suggests that exposure to innovation and education act as substitutes for one another.

The current evidence on the link between education and invention focuses on higher education (Toivanen and Väänänen 2016; Bianchi and Giorcelli 2020). This focus is natural since having post-secondary education is strongly correlated with being an inventor, even for our birth cohorts. But we know little about whether increases in primary and early secondary educational attainment have the potential to increase the likelihood an individual becomes an inventor. There are compelling indirect arguments on both sides. Bell et al. 2019 find a strong link between exceptional performance on an individual's standardized third grade math test and the likelihood of invention. In a similar spirit, Aghion et al. 2017 find a strong link between IQ score and inventor status. These facts argue in favor of the view that increasing educational attainment may not be useful in creating new inventors since the likelihood of invention may depend on innate qualities that cannot be altered simply by increasing an individual's years of schooling. On the other hand, there is a large literature documenting that the early years of life are very formative and have long run effects. Chetty et al. 2011 finds long-run effects from kindergarten classroom size and Barr and Gibbs 2022 find intergenerational effects from the early childhood program, Head Start. Likewise, Bell et al. 2019 find that early life exposure to local innovators makes individuals more likely to innovate later in life. These studies suggest that even relatively small shocks in early life can change important long-run outcomes such as income, health, and innovation. Since the compulsory schooling laws we utilize were focused on increasing education for children up to age 14, we have a setting that allows us to address whether primary and secondary education can form inventors.

2 Data

2.1 Matching the 1940 Census with Patents

Our goal is to construct a dataset that captures the lifetime innovative output of an individual, their educational attainment, and their exposure to compulsory schooling laws. To achieve this we use two main data sources. The first is the 1940 Full Count U.S. Census. The 1940 Census is chosen because it is the first U.S. Census to record the number of years of education an individual has attained. In addition, the 1940 Census provides an individual's state of birth and year of birth, which allows us to link to data on the compulsory schooling requirements an individual faced (Clay et al. 2021).

The second dataset comes from the PATSTAT database which provides us with comprehensive data on the worldwide patent record from 1899-2015. To link the 1940 Census with the patent record we match names on patents, as recorded in PATSTAT, with names from the 1940 Census. We drop all Census individuals where either the first or last name is a single character. We keep all inventor names on patents that successfully match with the HistPat database (Petralia et al. 2016). The HistPat database provides the location of patents granted by the United States Patent and Trademark Office (USPTO) from 1790-1975 and only includes patent × inventor observations where the inventor had residence in the U.S. at the time of application.

Before matching the list of names from the 1940 Census with the list of names from the patent record, we first process and clean both sets of names. In the PATSTAT data, we start by limiting to name strings that have two or three words, as this corresponds to the traditional naming convention of people having either a first and last name or a first, middle, and last name. Of the 1,211,782 name strings in the patent data, this restriction causes us to drop 33,547 name strings. Since inventor names are not recorded in a standardized fashion it is unclear which part of the name string is the first name, last name, and middle name or middle initial. As an example, US998585 was taken out by "MORRIS J LEAHY" while US705396 was taken out by "GARRISON GEORGE HENRY." To address this, we create all possible permutations of each name string.[1] Any match between a Census name and one of the permutations is considered to have matched the original PATSTAT name string. Table A1 provides an example of the permutations for the name string: "GARRISON GEORGE HENRY."

We then clean the names from the Census and PATSTAT using the Abramitzky, Boustan, and Eriksson (ABE) cleaner[2] which extracts the middle initial, substitutes in common male nicknames (i.e. Ben to Benjamin), and removes punctuation and occupational titles (Abramitzky et al. 2012; Abramitzky et al. 2014; Abramitzky et al. 2019). After the names are cleaned, we next remove any Census individuals who have a non-unique first name, last

[1]Middle names are shortened to be middle initials as only middle initials are recorded in the Census.
[2]The command is available as "abeclean" in STATA.

name, and middle initial. Limiting down to individuals with unique names in the 1940 Census helps to limit the number of false positive matches between the Census and the patent record. We also remove Census individuals who share their first and last name with at least one other Census individual who has a missing middle initial.[3] Removing these individuals from the Census ensures that all inventors match to one Census individual.

To match these uniquely named Census individuals to PATSTAT, we split the Census individuals into two mutually exclusive groups. The first group of Census individuals are those with a unique first and last name in the 1940 Census. For this group, we perform an exact match between the cleaned first and last names in the Census and patent records. We employ several criteria in order to reduce the incidence of false positive matches between PATSTAT and the Census. The second group of Census individuals are those who have a non-unique first and last name in the Census, but for whom all individuals with the same first and last name have non-missing middle initials.[4] For these individuals, we simply perform an exact match using their first name, last name, and middle initial. For a detailed discussion of the entire matching procedure, see Appendix A.

To examine whether these two groups of uniquely named Census individuals are similar to the overall population, we compare observable characteristics between the pool of uniquely named individuals who could potentially match to PATSTAT and the mutually exclusive group of individuals whom we remove from the pool of potential matches. Group (1) (first two columns) of Table 1 shows summary statistics from the 1940 Census for individuals who either have a unique first name, last name combination in the Census or have a unique first name, last name, middle initial combination and for whom none of the individuals who share a first name and last name have a missing middle initial. This comprises 47.8 million individuals or approximately 36% of all individuals in the Census and is the group of individuals that could potentially match with the PATSTAT data using our procedure. Group (2) of the table shows the same summary statistics but for the complement of the uniquely named group. The final column tests the difference in means between the groups.

[3]For example, the names "George H Garrison" and "George Garrison" would fit this criteria. In this instance we do not know the middle initial of the second "George Garrison" and so any inventor with the name "George Garrison" or "George H Garrison"could match to both "George H Garrison" and "George Garrison" in the Census.

[4]Census individuals who have duplicate first name, last name, and middle initial are removed.

While in all cases the differences between the groups are statistically significant, this is not surprising given the sample size, and the magnitude of the differences is generally small. On average, individuals in both groups have approximately 7.2 years of formal education and were born in 1909. Further, the characteristics of individuals' birth states when they were eighteen years old are similar across groups. The states of both groups have, on average, about 0.3 patents per thousand individuals, 45% of the population living in urban areas, and about $400 in manufacturing output per capita. The groups do display differences in terms of foreign born proportions. 11% of individuals who could match to the patent record are foreign born, whereas only 8% for the other group. This is natural since foreign names are more likely to be unique. Overall, Table 1 indicates that based on observable characteristics relevant to this study the group of uniquely named individuals is similar to those who are not uniquely named. This gives us more confidence that our results have external validity and may apply to the nearly two-thirds of individuals who could not match to the patent record.

To assess the quality of the match between PATSTAT and the Census, we use an inventor's implied age, based on Census data, and assess whether the implied age is within an individual's working years of life. If the implied age of the inventor at the date of application is either below eighteen or above seventy years old, then we consider the match to be invalid, since very few inventions are patented by individuals outside of these years. Kaltenberg et al. 2021 collect data on 1.5 million inventor ages associated with patents granted between 1976 and 2018. Their data show that only 2.8% of all inventor ages fall outside the ages of 18-70. Given that we are using historical data, it is likely that in our data the true share of inventions granted to inventors over the age of 70 is smaller since life expectancy was shorter in 1940 relative to the 1976-2018 time period. Later we will remove all matches where the implied age of the inventor is less than eighteen or greater than seventy, but for now we will keep these matches in order to assess the quality of our match.

As in Clay et al. 2021, we limit the Census data to all male individuals belonging to the 1885-1912 birth cohorts. We limit to males, as men accounted for most of the patenting at that time, and they don't experience name changes due to marriage which would challenge our match between the Census and patent record. Restricting to the 1885-1912 birth cohorts

Table 1: Comparison Between Uniquely Named Individuals and All Others

	Potential Match = 1		Potential Match = 0		Difference
	(1)		(2)		(3)
	Mean	St. Dev.	Mean	St. Dev.	Diff
Years of School	7.22	4.31	7.24	4.21	-0.02***
White	0.92	0.28	0.89	0.32	0.03***
Foreign Born	0.11	0.32	0.08	0.27	0.03***
Birthyear	1909.29	20.12	1908.78	20.20	0.51***
$\mathbb{1}\{\text{Compulsory Schooling} \geq 7\}$	0.84	0.37	0.81	0.39	0.03***
Patents per capita	0.31	0.20	0.30	0.20	0.01***
Urban share	0.45	0.22	0.44	0.22	0.01***
Mfg. Output per capita	401.01	282.72	391.52	282.10	9.49***
Observations	47,811,545		84,593,221		132,404,766

Notes: This table presents summary statistics on all individuals in the 1940 U.S. Full Count Census, with the individuals being split into two mutually exclusive and exhaustive groups. The first group, labeled as "Potential Match = 1" are individuals who have unique names as defined in Section 2 and could potentially match with PATSTAT. The second group, labeled as "Potential Match = 0" is the complement of this group. $\mathbb{1}\{\text{Compulsory Schooling} \geq 7\}$ is an indicator for whether the person is exposed to compulsory schooling laws which require them to complete at least 7 years of formal education. Patents per capita (thousand individuals), urban share, and mfg output to population ratio are measured from an individual's state of birth in the year the individual turns 18. $^{*}(p < 0.1)$,** $(p < 0.05)$,*** $(p < 0.01)$.

is intended to limit concerns about selection due to mortality as these individuals would be 27-54 at the time of the 1940 Census. In addition, the 1885-1912 birth cohorts experienced the first wave of compulsory schooling laws targeted at children up to age 14. Cohorts born before 1885 were born before the rollout of compulsory laws and cohorts born after 1912 experienced little variation in compulsory schooling exposure.

Table 2 provides summary statistics on the implied age of inventor × patent observations as of the application date for all inventors that match to this restricted set of individuals in the 1940 Census. Note that there are 512,882 patents that match to 102,438 uniquely named individuals in the 1940 Census. The average age of an inventor is 47, four years older than the average age of 43 in the Kaltenberg et al. 2021 data.[5] Just under 30% of all matches occur with an inventor that has an implied age outside the 18-70 window. If we take the average share of invalid matches for each inventor and then take the average across inventors, only counting each inventor once, we get that across inventors the average invalid proportion is 25%.

Figure 1 displays a histogram of the proportion of matches that are invalid for each inventor. The figure clearly shows that most individuals have either all valid matches or all invalid matches. To better understand the occurrence of individuals with all invalid matches, consider the case where an inventor's name matches to a uniquely named individual in the 1940 Census who did not patent, but someone with the same name born in a different generation did patent. Since all invalid matches will ultimately be removed, these individuals who have all invalid matches will be correctly marked as not being inventors when we perform our analysis. Figure 2 shows the distribution of implied ages across matches. The figure shows that most matches occur in middle age, consistent with Jaffe et al. 2021 who find that productivity for inventors peaks around their early 40s.

Overall, the evidence presented shows that the solid majority of matches in the data occur at appropriate ages and that the invalid matches are concentrated among individuals for whom all of their matches are invalid, and will therefore ultimately be removed from the sample before analysis. Since we only attempt to match to uniquely named individuals in

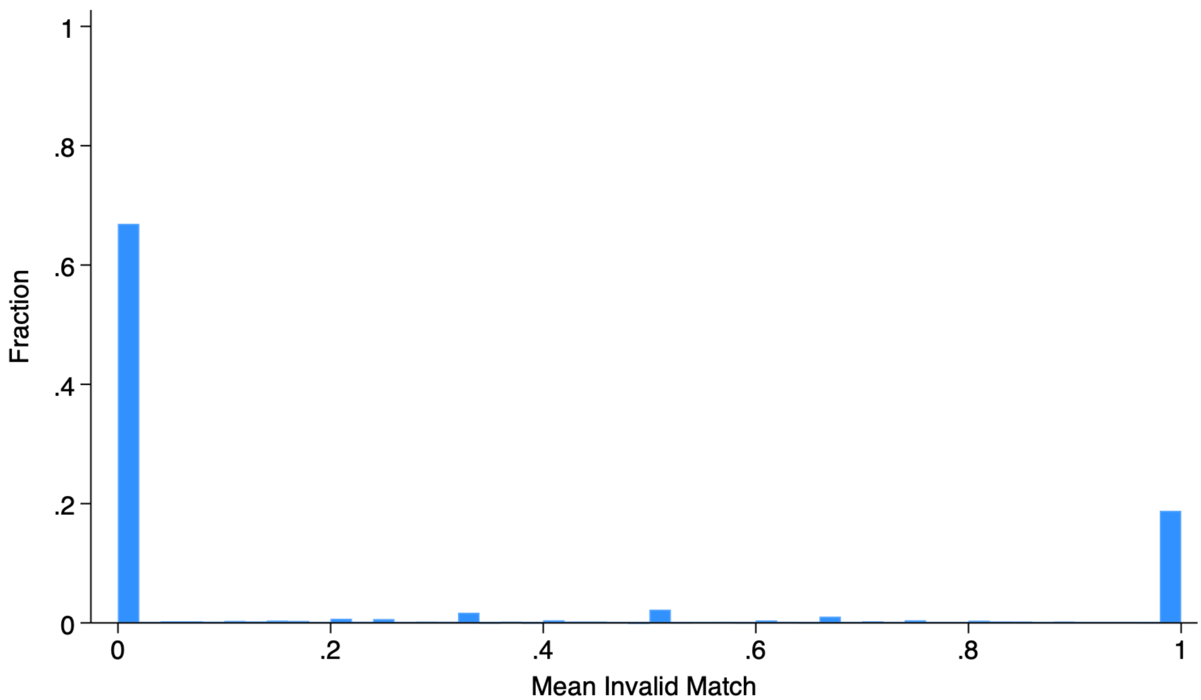
[5] Table A3 shows that if we only consider the subset of individuals who have no matches with implied ages at filing outside the 18-70 window, then the average age is 43.

Table 2: Match Summary Statistics

	Mean	SD	N
Age at Filing	47	27	512,882
Invalid Match	.29	.45	512,882
$\overline{\text{Invalid Match}}$	.25	.4	102,438
Matches	5	12	102,438

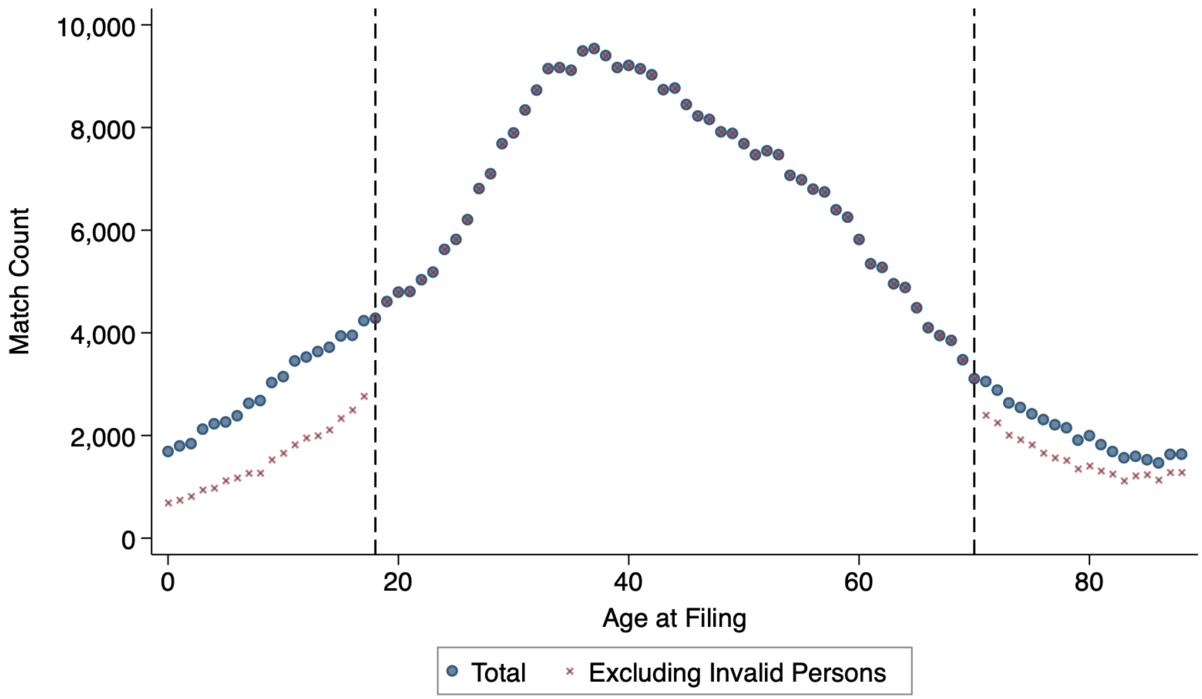
Notes: This table presents summary statistics on the implied age of inventor × patent observations that match to the 1940 Census, as well as summary statistics at the individual level. “Age at Filing” is defined as the year of application for the patent minus the individual's birth year. “Invalid Match” is defined as a match where the “Age at Filing” is either less than eighteen or greater than seventy. “$\overline{\text{Invalid Match}}$” is defined as $\frac{1}{J}\sum_{j=1}^{J}\{\frac{1}{N}\sum_{n=1}^{N}\mathbb{1}\{\text{Invalid Match}_n\}\}_j$ where j indexes each individual who has at least one match with the patent record and n indexes that individual's patents. “Matches” is the number of matches an individual has with the patent record.

Figure 1: Histogram of Invalid Match Shares Across Individuals



Notes: This figure displays a histogram of the share of matches that an individual has which are invalid matches where invalid matches are matches where the inventor has an implied age less than eighteen or greater than seventy.

Figure 2: Distribution of Implied Ages Across Matches



Notes: This figure shows the number of matches which occur at every age between 0-88 for two samples. The first plot, in blue circles, is all matches. The second plot, in red Xs, is all matches excluding persons who only have invalid matches. Vertical lines are placed at 18 and 70, the thresholds for a valid age on a patent.

the Census, we are highly confident in the validity of the matches that we do obtain where the implied age at patenting is appropriate. Finally, the distribution of implied ages in the data matches nicely with what we know about the lifecycle of an inventor (Jaffe et al. 2021). After matching Census individuals to the patent record, we are able to obtain information on their lifetime patenting activity, such as the number of patents they were granted. We winsorize patenting activity at the 1st and 99th percentile of the non-zero distribution in order to mitigate the impact of outliers.[6] In addition, the 1940 Census directly provides the number of years of education that an individual achieves.

2.2 Compulsory Schooling Laws

To measure the amount schooling each individual was required to have due to compulsory schooling laws we rely on Clay et al. 2021 who provide a newly collected dataset on compulsory schooling laws for the 1880-1930 birth cohorts. The data were constructed using the session laws of individual states and carefully take into account child labor and continuation schooling laws which often interacted with compulsory schooling laws to determine the required amount of schooling an individual must obtain.[7] The calculation assumes that children entered school at the required age, attended each year they were required (including continuation school), and left for the workforce as soon as legally possible. The required amount of schooling was often enforced through fines. Teachers and police were tasked with enforcement, although there was variation in how the laws were enforced. For example, Michigan's 1871 compulsory schooling law states that: "In case any parent, guardian, or other person shall fail to comply with the provisions of this act, said parent, guardian, or other person shall be liable to a fine of not less than five dollars or more than ten dollars for the first offense, nor less than ten or more than twenty dollars for the second and every subsequent offense." We link compulsory schooling laws to individuals based on their state and year of birth.[8]

[6] This effectively caps the maximum number of patents an individual can obtain in their lifetime at 34 and the number of lifetime breakthrough innovations at 10.

[7] Continuation schooling was designed to provide an education for children who opted to leave regular schooling to work.

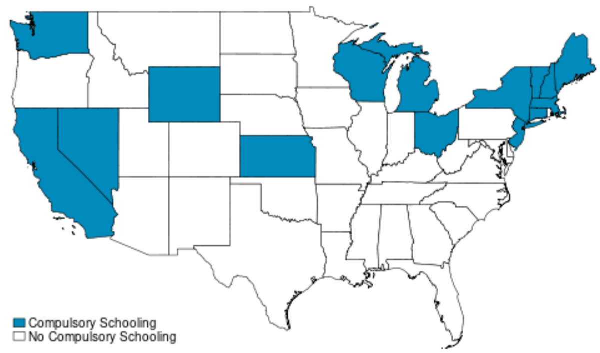
[8] While state of birth will not always correspond to the state in which the went to primary and secondary school, it is the best source of information on the topic that is provided in the Census.

Figure 3 shows the rollout of compulsory schooling laws by decade based on an individual's state and year of birth. In 1870, only Massachusetts, Vermont, and the District of Columbia had adopted compulsory schooling laws. Southern states were slow to adopt the law as the Republican party strongly advocated for compulsory schooling laws (McAfee 1998). Figure 3 shows that by 1900 nearly every state outside the South had some sort of compulsory schooling law adopted, while no Southern state had adopted. What sparked this flurry of compulsory schooling laws? Advocates of compulsory schooling during this time were concerned with having universal suffrage without an educated population, the relationship between crime and illiteracy, and the assimilation of immigrants (Eisenberg 1988; Bandiera et al. 2019). By 1910, the wave of compulsory schooling requirements had even extended to the Southern states, and by 1920 every state had adopted a compulsory schooling law.

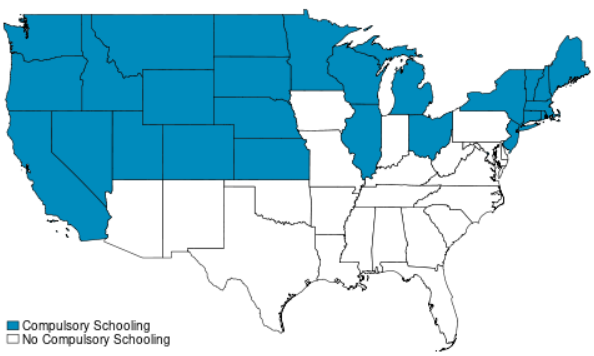
Given the disparities in adoption between the South and other parts of the country, our empirical analysis will include careful treatment of this regional heterogeneity in compulsory schooling adoption. As Stephens and Yang 2014 shows during a later wave of compulsory schooling laws, it is plausible that much of the contemporaneous shocks that are correlated across space and time with compulsory schooling changes happen at the region-year, rather than the state-year, level. Thus, following Stephens and Yang 2014 and Clay et al. 2021, we control for region-year fixed effects in all specifications in order to ensure that all comparisons are being made within region. Another concern with using compulsory schooling laws to obtain variation in educational attainment is that the increases in educational attainment caused the adoption of compulsory schooling laws themselves. This possible reverse causality is the heart of the critique found in Landes and Solmon (1972) and Lleras-Muney (2002). To address this concern, we show that compulsory schooling laws only increased educational attainment when the requirements reached a high enough threshold to bind, which was often several years after adoption of the law. If increases in schooling were causing states to adopt the laws then we would expect to see effects even as the requirements of the law were being phased in. We find no evidence of this. In addition, qualitative evidence supports the idea that compulsory schooling laws increased school attendance. Public administrators of Kentucky, Indiana, Missouri, and North Carolina all stated that compulsory schooling

Figure 3: Compulsory Schooling Law Adoption

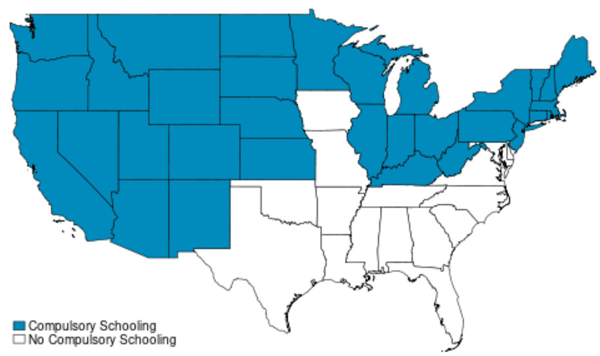
(a) Year 1870



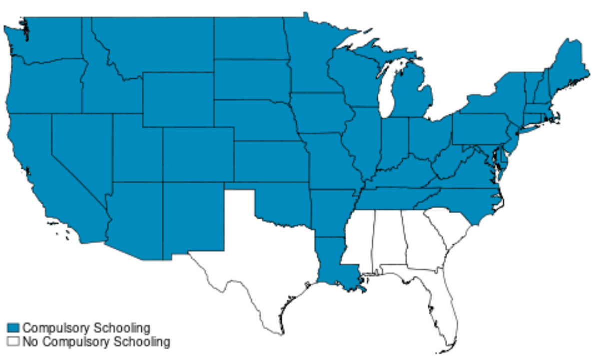
(b) Year 1880



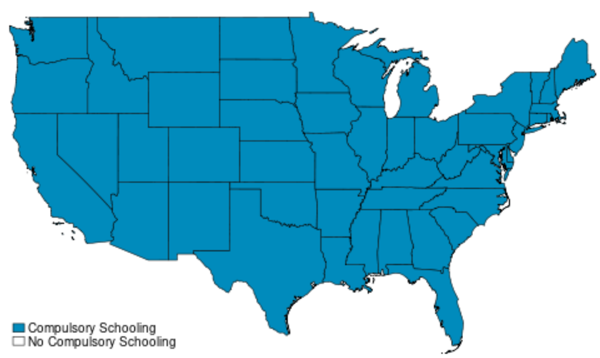
(c) Year 1890



(d) Year 1900



(e) Year 1910



(f) Year 1920

Notes: The figure shows the decade in which each U.S. state enacts their first compulsory schooling law by highlighting each state blue when they enact their first compulsory schooling law.

laws increased attendance (Lingwall 2014). For example, Indiana administrators stated that their state's laws: "have succeeded in placing many children in the Indianapolis schools who probably would otherwise not have been there..."[9]

2.3 Local Characteristics

We collect data on a state's population, share of the population living in urban areas, and manufacturing output per capita from the 1880, 1890, 1900, 1920, and 1930 decennial censuses (Haines et al. 2010). We linearly interpolate these variables to arrive at estimated measures in each state × year observation from 1880-1930. Since the PATSTAT data doesn't geographically locate patents, and it only goes back to 1899, we utilize the HistPat database of Petralia et al. 2016 to measure the number of patents applied for in each state × year observation from 1880-1930. By combining this with our population data, we measure local innovation intensity as state patenting per thousand individuals. We match each individual to these variables using an individual's state of birth and the year the individual turned eighteen. We use these variables as controls in some specifications and later interact them with variation in compulsory schooling laws to explore heterogeneity and mechanisms driving our main effect.

2.4 Descriptive Statistics

Table 3 displays descriptive statistics on our sample of uniquely named men belonging to the 1885-1912 birth cohorts. The individuals are split by whether they are an inventor or not where an inventor is defined as an individual that applies for one granted patent in their lifetime, as measured by our match to PATSTAT. A little over 1% of our sample are inventors, almost all of which are white. On average, inventors have about 1.5 more years of education than non-inventors. While the correlation between educational attainment and invention shows an association between education and invention, our empirical strategy will use variation in compulsory schooling laws to estimate the causal impact of education on the formation of inventors. Inventors grew up in states that were more urban and had higher

[9] Indiana Department of Public Instruction (1901, 508).

Table 3: Descriptive Statistics by Inventor Status

	Inventor = 1		Inventor = 0		Difference
	(1)		(2)		(3)
	Mean	St. Dev.	Mean	St. Dev.	Diff
Years of School	10.57	3.73	8.92	3.53	1.65***
White	0.97	0.17	0.92	0.26	0.05***
Birthyear	1898.88	7.96	1900.34	7.94	-1.46***
$\mathbb{1}\{CS \geq 7\}$	0.59	0.49	0.56	0.50	0.03***
Patents	3.11	5.02	0.00	0.00	3.11***
Breakthrough Innovations (Top 20%)	0.54	1.44	0.00	0.00	0.54***
Patents per capita	0.35	0.19	0.30	0.20	0.05***
Urban share	0.51	0.21	0.47	0.22	0.04***
Mfg. Output per capita	497.69	268.75	466.28	277.45	31.42***
Observations	83,217		6,604,037		6,687,254

Notes: This table presents summary statistics on all uniquely named men from the 1885-1912 birth cohorts in the 1940 Census. The individuals are broken out by whether they ever match with the patent record (Inventor = 1) or not (Inventor = 0). $\mathbb{1}\{\text{Compulsory Schooling} \geq 7\}$ is an indicator for whether the person is exposed to compulsory schooling laws which require them to complete at least 7 years of formal education. Patents per capita (thousand individuals), urban share, and mfg output to population ratio are measured from an individual's state of birth in the year the individual turns 18. *(p<0.1), **(p<0.05), ***(p<0.01).

patenting activity and manufacturing productivity, consistent with Bell et al. 2019 who find, in a modern context, that exposure to innovation at young ages is crucial to the formation of inventors.

3 Empirical Strategy & Results

3.1 Poisson Estimation

Our empirical strategy seeks to identify the causal effect that formal education has on an individual's lifetime innovative output. Plausibly exogenous variation in the years of schooling an individual completes comes from the rollout of compulsory schooling laws across the U.S. which caused some children to remain in school longer than they otherwise would have. To implement this strategy we start by estimating the following Poisson model via quasi-maximum likelihood:

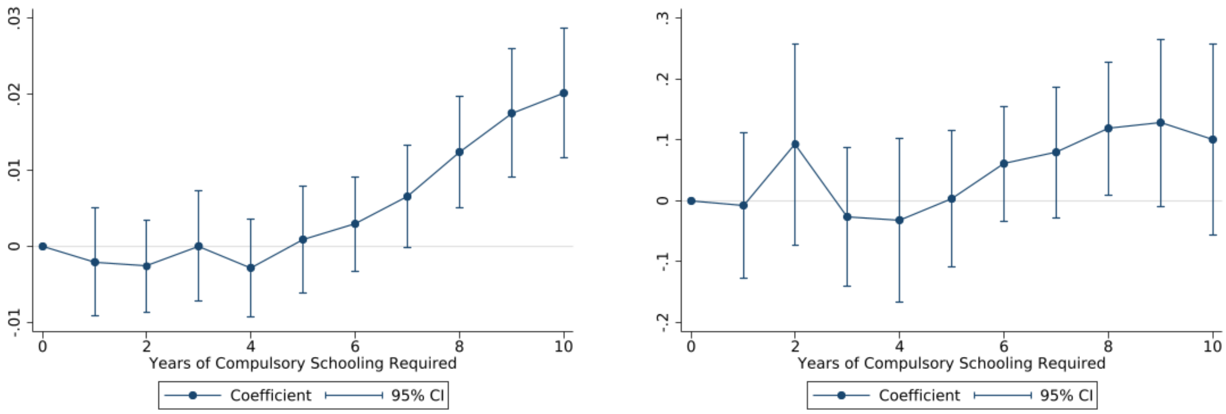
$$E[\mathrm{Y}_{icsr}|\mathrm{X}_{icsr}] = exp\left(\sum_{k=1}^{10}\beta_k * \mathbb{1}\{\mathrm{CS}_{sc}=k\} + \zeta_s + \tau_{cr} + \Psi\mathrm{X}_{ics}\right) \quad (1)$$

The dependent variable, Y_{icsr} is a count variable such as the number of years of schooling that individual i, part of birth cohort c, born in state s, and Census region r, attained, or the individual's lifetime number of granted patents. This is regressed on dummy variables for the various amounts of compulsory schooling required of an individual. Based on an individual's state and year of birth, that person can be required to complete anywhere between zero to ten years of schooling, which is recorded in the variable CS_{sc}. Having zero years of compulsory schooling required is the omitted category. Indicators for the various amounts of compulsory schooling required are preferred to using numeric values as the dummy variables allow for non-linearities in the effect of compulsory schooling. Indeed, our results will exhibit non-linearities, as we will show that moving from 4 to 5 years of compulsory schooling produces no change in the years of schooling attained for an individual, whereas moving from 8 to 9 years of schooling required generates an increase in educational attainment. State fixed effects removes time-invariant state level heterogeneity at the state level. The inclusion of region by birth-cohort fixed effects prevents us from conflating regional heterogeneity with the effect of compulsory schooling laws by making comparisons within region. Our identification assumption is that without compulsory schooling laws, educational outcomes would have evolved similarly for states in the same region. X_{ics} is a vector of controls which includes an indicator for whether the individual belongs to a minority race;[10] and the number of patents per capita, urban share, and manufacturing output per capita in the individual's state of birth, measured in the year the individual turns eighteen. We cluster standard errors at the state $\times$ year of birth level as in Clay et al. 2021.

Figure 4 presents the resulting β_k coefficients, along with their 95% confidence intervals, from estimating Equation 1 with various dependent variables. In Panel (a) of Figure 4 the dependent variable is a count of the years of schooling the individual completed. When one to five years of compulsory schooling are required, the effect on educational achievement is

[10]Minority races are all races other than "White."

Figure 4: Effect of Compulsory Schooling on Education and Invention



(a) Years of School

(b) Patents

Notes: This figure presents the coefficients and 95% confidence intervals, from estimating Equation 1. The vector of controls X_{ics} includes a dummy variable for whether the individual is of a minority race. Standard errors are clustered at the state × year of birth level. Panel (a) shows the results when the dependent variable is a count of the number of years of school an individual achieves. In panel (b) the dependent variable is the number of lifetime patents the individual attains.

indistinguishable from having zero years of compulsory schooling required. This is reassuring, as it seems unlikely that requiring less than six to seven years of education would have a binding effect on most individuals. In addition, if states were adopting compulsory schooling laws because they were seeing increases in educational attainment, we would expect to see positive coefficients on these low levels of required education. Reassuringly, the positive effect of compulsory schooling requirements emerge only after the requirements are sufficiently high. The coefficient increases when six years of schooling are required, although it remains statistically insignificant. When seven years of schooling are required, there is a statistically significant effect on schooling achievement. The effect increases in a monotonic fashion, with ten years of required schooling increasing educational attainment by around 2%, in expectation.

Panel (b) of Figure 4 presents the same results but with the count of lifetime patents as the dependent variable. Similar to when years of schooling was the dependent variable, there is no effect of compulsory schooling on patenting activity when five or fewer years of school are required. Once six years of schooling are required, the coefficients start to increase in a nearly monotonic fashion, peaking at around a 10% increase in patenting when eight to ten years of schooling are required.

Table 4 presents the results of estimating a similar specification to Equation 1 but where a dummy for zero to five years of compulsory schooling is the omitted category, and nine or ten years of compulsory schooling required are collapsed to a single dummy (since very few states required ten years of education). In column (1), with only the minority race control, we see an increasing effect of compulsory schooling requirements on educational attainment, with the effect peaking at nine or more years of schooling required, which leads to about a 2% increase in years of schooling achieved. In column (2) controlling for characteristics of an individual's childhood state doesn't affect the coefficients of interest, although we do see that increased patenting and manufacturing productivity are positively correlated with educational attainment. In columns (3) and (4), the count of lifetime patents is the dependent variable, and we see that requiring nine or more years of compulsory schooling increases patenting counts by over 10% regardless of the controls used.

Not all inventors are equally prolific (Akcigit et al. 2016), and not all inventions are equally valuable (Kogan et al. 2017). Indeed, both distributions are right skewed with a few inventors and inventions producing a significant amount of the total value. In columns (5) and (6) we test whether the patents produced from the additional education required are impactful patents, since it could be the case that compulsory schooling led to patents of limited usefulness. To do this, we determine whether each patent is a breakthrough innovation using the text-based measure created in Kelly et al. 2021.[11] Kelly et al. 2021 label patents as breakthrough innovations if the patent displays significant textual dissimilarity to the text of patents that came before it and high amounts of textual similarity to patents that come after it. The intuition is that patents which are highly original (dissimilar to the patents that came before it) and impactful in spurring innovation (similar to the patents that come after it) are breakthrough innovations.[12] The pattern and magnitude of the coefficients is similar to what we observe when lifetime patent counts are the dependent variable. This suggests that the innovations generated by compulsory schooling requirements had the same

[11] The number of citations a patent receives would be another way to measure a patent's value, but citations are consistently recorded by the USPTO in patent documents only after 1947 (Kelly et al. 2021).

[12] Specifically, we denote patents that are above the 80th percentile in the Kelly et al. 2021 5 year window reverse-forward similarity metric as breakthrough innovations. All other patents are not breakthrough innovations.

Table 4: The Effect of Compulsory Schooling (Poisson)

	Years of School		Patents		Breakthrough Innovs.	
	(1) Poisson	(2) Poisson	(3) Poisson	(4) Poisson	(5) Poisson	(6) Poisson
$\mathbb{1}\{\text{CS} = 6\}$	0.004**	0.004**	0.062**	0.061*	0.067	0.072
	(0.002)	(0.002)	(0.032)	(0.032)	(0.044)	(0.045)
$\mathbb{1}\{\text{CS} = 7\}$	0.007***	0.008***	0.079**	0.075*	0.093*	0.100*
	(0.002)	(0.002)	(0.039)	(0.040)	(0.056)	(0.058)
$\mathbb{1}\{\text{CS} = 8\}$	0.013***	0.013***	0.120***	0.113***	0.125**	0.138**
	(0.002)	(0.002)	(0.039)	(0.041)	(0.057)	(0.060)
$\mathbb{1}\{\text{CS} \geq 9\}$	0.019***	0.017***	0.122**	0.112**	0.126*	0.147*
	(0.003)	(0.003)	(0.053)	(0.056)	(0.075)	(0.080)
Minority Race	-0.408***	-0.408***	-0.700***	-0.700***	-0.679***	-0.679***
	(0.007)	(0.007)	(0.042)	(0.042)	(0.064)	(0.064)
Patents per capita		0.010***		0.006		0.055
		(0.002)		(0.038)		(0.054)
Urban share		0.006		-0.059		-0.157
		(0.007)		(0.116)		(0.174)
Mfg. Output per capita		0.012***		0.087		-0.042
		(0.004)		(0.081)		(0.118)
$\overline{\text{Y}}$	8.94	8.94	0.04	0.04	0.01	0.01
Observations	6,687,254	6,687,254	6,687,254	6,687,254	6,687,254	6,687,254

Notes: This table presents results from estimating equation (1) using Poisson pseudo-maximum likelihood. "Years of School" is a count of the number of years of formal schooling the individual achieved. "Patents" is a count of the lifetime number of patents granted to the individual. "Breakthrough Innovs." is a count of the lifetime number of breakthrough patents granted to the individual, which are patents which belong to the top 80 percentile of reverse-forward similarity in a five year window, as measured in Kelly et al. 2021. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth × year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).

propensity to be breakthrough innovations as what would be expected from a random sample of patents.

3.2 Two Stage Least Squares Estimation

The results in the previous section clearly show that compulsory schooling led to an increase in educational attainment when individuals were exposed to sufficiently high requirements. In addition, these same individuals went on to patent more. The innovations generated from the additional schooling were not limited to incremental innovations but included a significant number of breakthrough innovations. Looking at Table 4, we see that the requirement of nine or more years of schooling increased educational attainment by 2% and increased innovative activity by around 12%. This suggests that in our context the elasticity of innovation with respect to formal education is approximately six, although this number will change as different amounts of compulsory schooling requirements are chosen.

As an alternative approach to estimating the elasticity of innovation with respect to education, we now employ a 2SLS estimation procedure where variation in compulsory schooling exposure serves as an instrumental variable (IV) for educational attainment. This will allow us to summarize the magnitude and precision of the relationship between formal education and invention in one coefficient, and it will allow us to test the strength of the relationship between compulsory schooling and educational attainment using a conventional F-statistic test.

To implement this 2SLS procedure we estimate the following specification where $\mathrm{K} \equiv \{6, 7, 8, \geq 9\}$ as in Clay et al. 2021:

$$\text{Years of School}_{icsr} = \sum_{k \in K} \beta_k * \mathbb{1}\{\mathrm{CS}_{sc} = k\} + \zeta_s + \tau_{cr} + \mathrm{X}_{ics} + \upsilon_{icsr} \tag{2}$$

$$\mathrm{Y}_{icsr} = \beta \widehat{\text{Years of School}}_{icsr} + \gamma_s + \delta_{cr} + \mathrm{X}_{ics} + \varepsilon_{icsr} \tag{3}$$

In our preferred specification, we use a binary variable indicating whether an individual ever patented in their lifetime as the dependent variable.[13] Table 5 displays the results with columns (1) and (2) estimating Equation 3 using Ordinary Least Squares (OLS) where an individual's endogenous years of schooling is used instead of the predicted value from

[13] For completeness, Table B1 and Table B2 respectively display the results when the count of lifetime patents and count of breakthrough innovations are the dependent variables, but they are not included in the main body of the paper as count variables are more appropriately modeled using Poisson estimation.

Equation 2. Estimating via OLS is likely to result in biased estimates of the effect of formal education on the likelihood of becoming an inventor since schooling is an endogenous choice and is correlated with many omitted variables. For example, creative and motivated people may be more likely to achieve higher levels of education, which would bias the OLS estimates upwards. On the other hand, many studies estimating the effect of education on a myriad of outcomes find that IV estimation yields larger effects of education, plausibly because the local average treatment effect diverges from the average treatment effect (Card 1999). In columns (1) and (2), there is a strong positive association between educational attainment and the likelihood an individual becomes an inventor. An additional year of schooling is associated with an 0.16 percentage point increase in the likelihood of being an inventor, regardless of the control variables used. To put the estimate in perspective, the mean educational attainment in the sample is 8.9 years of schooling, so an additional year of schooling increases formal education by 11%. This 11% increase in formal education is associated with a 13% increase off the mean probability of becoming an inventor of 1.2%.[14]

In columns (3) and (4) we estimate Equation 3 via 2SLS. In both columns the first stage is strong with Kleibergen-Paap F-statistics that are at or above the conventional threshold of ten.[15] In column (3), when a dummy for being a "Minority Race" is the only control variable, the effect of formal education on becoming an inventor doubles relative to the OLS specifications, and we estimate that an additional year of schooling increases the likelihood someone becomes an inventor by 0.39 percentage points. This is a 31% increase off the mean, suggesting that the elasticity of becoming an inventor with respect to formal education is approximately three.[16] When controls for local innovative and economic conditions are included, the estimate increases so that an additional year of schooling generates a 45% increase off the mean probability of becoming an inventor. Recall that our back-of-the-envelope estimates of elasticities with the Poisson model were around six when nine or more years of compulsory schooling were required. This is consistent with the elasticities of 3-5 that we are recovering in the 2SLS estimation.

In Table 6 we examine whether additional schooling has an effect on the likelihood an

[14]This is calculated as $\frac{0.161}{1.244}$

[15]The corresponding first stage estimation results are available in Table B3.

[16]This is calculated as $\frac{31}{11}$

Table 5: The Effect of Compulsory Schooling on Becoming an Inventor (2SLS)

	$\mathbb{1}\{\text{Inventor}\} \times 100$			
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Years of School	0.161***	0.161***	0.388**	0.562**
	(0.003)	(0.003)	(0.192)	(0.257)
Minority Race	-0.001	-0.001	0.628	1.112
	(0.016)	(0.016)	(0.532)	(0.712)
Patents per capita		-0.035		-0.059
		(0.032)		(0.036)
Urban share		-0.198**		-0.323***
		(0.078)		(0.116)
Mfg. Output per capita		0.109*		0.052
		(0.057)		(0.071)
$\overline{\text{Y}}$	1.244	1.244	1.244	1.244
KP *F*-Stat			17.07	9.64
Observations	6,687,254	6,687,254	6,687,254	6,687,254

Notes: This table presents results from estimating Equation 3 using 2SLS. "$\mathbb{1}\{\text{Inventor}\} \times 100$" is an indicator for whether the individual ever patented in their lifetime, multiplied by 100. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth × year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).

individual goes on to become a breakthrough inventor, as defined by ever patenting a breakthrough innovation according to the same criteria we used earlier. Relative to the effects on inventor status, we find similarly sized positive effects of education on the likelihood that an individual becomes a breakthrough inventor; these effects become precisely estimated with the use of local characteristics as controls in column (4). Overall, we find large causal effects of formal education on the extensive margin of becoming an inventor.

Table 6: The Effect of Compulsory Schooling on Becoming an Inventor (2SLS)

	$\mathbb{1}\{\text{Breakthrough Inventor}\} \times 100$			
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Years of School	0.058***	0.058***	0.070	0.188*
	(0.001)	(0.001)	(0.084)	(0.110)
Minority Race	0.045***	0.045***	0.079	0.406
	(0.007)	(0.007)	(0.232)	(0.305)
Patents per capita		0.009		0.002
		(0.014)		(0.016)
Urban share		-0.120***		-0.160***
		(0.035)		(0.050)
Mfg. Output per capita		0.025		0.006
		(0.025)		(0.030)
$\overline{\text{Y}}$	0.321	0.321	0.321	0.321
KP *F*-Stat			17.07	9.64
Observations	6,687,254	6,687,254	6,687,254	6,687,254

Notes: This table presents results from estimating Equation 3 using 2SLS. "$\mathbb{1}\{\text{Breakthrough Inventor}\} \times 100$" is an indicator for whether the individual ever patents a breakthrough innovation, defined as a patent which belongs to the top 80 percentile of reverse-forward similarity in a five year window, as measured in Kelly et al. 2021. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth $\times$ year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).

3.3 Heterogeneity in Effect by Local Conditions

Bell et al. 2019 find that exposure to innovation at young ages is an important factor in determining an individual's likelihood of becoming an inventor. This suggests that all else equal, individuals growing up in areas with more innovation occurring around them will be more likely to become inventors. Our results above demonstrate that exposure to compulsory education also increases invention. In light of both of these results, a natural question is: how do exposure to innovation and education interact to determine inventor formation and subsequent innovative output? One possibility is that growing up in an innovative geographic area is complementary to increases in educational attainment. After all, the formal content of secondary education is basic mathematics, basic science, grammar, history, and language; none of these subjects are covered to a sufficiently exhaustive degree to bring students to the frontier of invention, or to provide them with knowledge of what inventions are needed in their local industries. Therefore, it's plausible that exposure to innovation is a necessary condition for the creation of inventors; in that case, the causal effect of education on innovation should be close to zero if there is no exposure to innovation in the student's community. On the other hand, exposure to innovation and education could act as substitutes for one another in the creation of inventors: those who have little exposure to innovation in childhood may have the most to gain from increased education.

To distinguish between these competing stories, we augment our empirical strategy by fully interacting the endogenous and instrumental variables with various characteristics capturing the innovative and economic activity the individual was exposed to during their early years of life. These variables are each measured in the individual's state of birth in the year the individual turned eighteen. We use the number of patents granted per capita to measure local innovation intensity. Individuals living in states with high per capita patenting are more likely to be exposed to innovation at younger ages and have greater opportunities to innovate when they get older. Manufacturing output per capita measures economic activity and the share of the population living in urban areas captures general economic development, as well as opportunities for agglomeration economies. These three variables are standardized to have mean zero and standard deviation of one in each year. De-meaning the variables prevents conflation from time trends and standardizing the variable allows for ease of comparison across the three measures. A positive interaction term between these measures and educational attainment would provide evidence that the effect of education on invention was stronger in states that had high innovation intensities. This would suggest that education complements exposure to innovation so that additional education is most beneficial in forming inventors when individuals are also living in innovative areas. On the other hand, a negative coefficient would suggest that education and local innovation intensity are substitutes.

Table 7 displays the results when an indicator for whether the individual ever patents is the dependent variable. In column (1), we fully interact years of schooling attained with state patenting per capita. We find a strong, negative interaction term between local innovative activity and educational attainment. The effect that an additional year of schooling has on the probability that an individual becomes an inventor is approximately $\frac{-.062}{.485} = 13\%$ lower for an individual who is one standard deviation above the mean state patenting per capita measure. To put this in perspective, consider two individuals who both turned 18 in 1910, but one was born in Alabama and the other in California.[17] On average, the individual living in Alabama would see their probability of becoming an inventor increase by 0.485 + (-1.13 $\times -0.062) = 0.56$ percentage points when given one extra year of schooling while the person in California would only have a marginal effect of $0.485 + (1.67 \times -0.062) = 0.38$ percentage points. This is after controlling for the substantial direct effect that state patenting per capita has on an individual's likelihood of becoming an inventor as seen by the highly significant and positive coefficient on state patents per capita. The evidence suggests that local innovative activity and educational attainment are substitutes in the formation of inventors.

While patenting per capita is a direct measure of innovative intensity, we next use measures that capture broader notions of economic activity by interacting the urban share and manufacturing output per capita with educational attainment in columns (2) and (3) respectively. We find similar results, indicating that the results are robust to replacing local innovative activity with the similar but related concept of economic activity. When all three

[17] Alabama is 1.13 standard deviations below the mean state patenting per capita measure in 1910 while California is 1.67 standard deviations above the mean.

Table 7: Heterogeneity by Local Conditions in the Formation of Inventors

	$\mathbb{1}\{\text{Inventor}\} \times 100$			
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Years of School	0.485**	0.464**	0.506***	0.529***
	(0.218)	(0.193)	(0.194)	(0.137)
Years of School × Patents per capita	-0.062***			-0.134***
	(0.018)			(0.043)
Years of School × Urban share		-0.052**		0.089
		(0.023)		(0.057)
Years of School × Mfg. Output per capita			-0.049*	0.020
			(0.028)	(0.046)
Patents per capita	0.558***	-0.033	-0.032	1.235***
	(0.189)	(0.036)	(0.040)	(0.406)
Urban share	-0.157	0.273	-0.225*	-1.002*
	(0.122)	(0.284)	(0.124)	(0.512)
Mfg. Output per capita	0.069	0.086	0.569*	-0.170
	(0.068)	(0.067)	(0.309)	(0.473)
Minority Race	1.042*	0.942*	1.048**	1.126***
	(0.595)	(0.529)	(0.516)	(0.377)
$\overline{\text{Y}}$	1.24	1.24	1.24	1.24
KP *F*-Stat	7.17	8.62	7.67	7.08
Observations	6,687,254	6,687,254	6,687,254	6,687,254

Notes: This table presents results from estimating Equation 3 using 2SLS where compulsory schooling indicators and "Years of School" are interacted with the characteristic of interest (i.e. Patents per capita). "$\mathbb{1}\{\text{Inventor}\} \times 100$" is an indicator for whether the individual ever patented in their lifetime, multiplied by 100. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth × year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).

of these correlated measures are included in column (4), the negative effect loads on patenting per capita, the most direct measure of local innovative activity.

The results of Bell et al. 2019 suggest that exposing individuals to innovation at a young age will increase the likelihood they go on to become an inventor. One way to do this is to provide incentives for innovative activity to become more geographically dispersed. While this would decrease the intensity of exposure for some individuals, it would increase the breadth of exposure, which Bell et al. 2019 document as being particularly important. But policies of this nature would come at a substantial productivity loss. Indeed, Moretti 2021 finds that the clustering of innovative activity generates large productivity benefits for star inventors.

Our results suggest that increased educational attainment could play a similar role by disproportionately increasing the likelihood of invention for those who live in less innovative areas. Furthermore, this would lower inequality in opportunity as it would increase the probability of invention for those living in less economically developed areas. Overall, these results indicate that exposure to innovation and education act as substitutes in the formation of inventors. This finding suggests that economic development and local innovative intensity are not necessary conditions for education to be effective at creating inventors; on the contrary, they are substitutes for education.

4 Conclusion

This paper explores the importance of education in forming inventors in the context of the United States' first wave of compulsory schooling laws. To do so we retrieve the lifetime patenting activity of uniquely named individuals by matching the Full Count 1940 Census with the patent record found in PATSTAT. Using both Poisson and 2SLS models, where the compulsory schooling laws serve as instruments for educational attainment, we find that increases in primary and secondary schooling had a large effect on lifetime innovative output as measured by an individual's patenting activity. We find that a marginal year of schooling significantly increases the number of granted patents a person files for and the probability that an individual patents a breakthrough innovation. Further, we find that this positive effect of education is decreasing in the amount of local innovative and economic activity an individual is exposed to. This result suggests that, in the production of inventors, education and local innovative activity are substitutes.

Our results suggest that increasing formal education is a way to unlock the potential of "lost Einsteins." Our findings support this view in two respects. First, increased educational attainment raises the likelihood an individual becomes an inventor, even for those who are born into highly innovative and economically developed geographies. Second, education is particularly effective at creating inventors when individuals are least exposed to innovation and thus have much fewer opportunities to reach their potential of becoming an inventor. Education can help lower the geographic opportunity gap.

The historical setting of our study provides ample opportunity for future work to examine these issues in a modern setting. While the compulsory schooling laws used in our study increased educational attainment by requiring seven to ten years of schooling, understanding the effect of increasing years of schooling along the educational distribution in a modern context would be a valuable contribution. While moving individuals from seven to eight years of schooling was sufficient to increase the probability of a breakthrough innovation in this historical context, the increase in the burden of knowledge (Jones 2009), the decline of the independent inventor (Lamoreaux and Sokoloff 2005), and the increase in average educational attainment (Goldin 1998) ensure that more education is needed to form an inventor today than one hundred years ago. Further, this study does not address the quality of education. As high school and college graduation rates have increased over time, the policy discussion has shifted from simply requiring a certain level of education to increasing the value added of schooling. Understanding how specific policies aimed at increasing the quality of education impact the formation of inventors would provide a guide to policymakers and scholars seeking to understand the link between education and innovation. Overall, our results point to the importance of education in forming inventors.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson (2019). "To the new world and back again: Return migrants in the age of mass migration". In: *ILR Review* 72.2, pp. 300–322.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2012). "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration". In: *American Economic Review* 102.5, pp. 1832–1856.

— (2014). "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration". In: *Journal of Political Economy* 122.3, pp. 467–506.

Aghion, Philippe et al. (2017). *The social origins of inventors*. Tech. rep. National Bureau of Economic Research.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva (2016). "Taxation and the international mobility of inventors". In: *American Economic Review* 106.10, pp. 2930–81.

Bandiera, Oriana et al. (2019). "Nation-building through compulsory schooling during the age of mass migration". In: *The Economic Journal* 129.617, pp. 62–109.

Barr, Andrew and Chloe R Gibbs (2022). "Breaking the cycle? intergenerational effects of an antipoverty program in early childhood". In: *Journal of Political Economy* 130.12, pp. 3253–3285.

Bell, Alex et al. (2019). "Who becomes an inventor in America? The importance of exposure to innovation". In: *The Quarterly Journal of Economics* 134.2, pp. 647–713.

Bianchi, Nicola and Michela Giorcelli (2020). "Scientific Education and Innovation: From Technical Diplomas to University Stem Degrees". In: *Journal of the European Economic Association* 18.5, pp. 2608–2646. ISSN: 1542-4766. DOI: 10.1093/jeea/jvz049.

Card, David (1999). "The causal effect of education on earnings". In: *Handbook of labor economics*. Vol. 3. Elsevier, pp. 1801–1863.

Chetty, Raj et al. (2011). "How does your kindergarten classroom affect your earnings? Evidence from Project STAR". In: *The Quarterly journal of economics* 126.4, pp. 1593–1660.

Clay, Karen, Jeff Lingwall, and Melvin Stephens Jr (2021). “Laws, educational outcomes, and returns to schooling evidence from the first wave of US state compulsory attendance laws”. In: *Labour Economics* 68, p. 101935.

Eisenberg, Martin Jay (1988). “Compulsory attendance legislation in America, 1870 to 1915”. In.

Goldin, Claudia (1998). “America’s graduation from high school: The evolution and spread of secondary schooling in the twentieth century”. In: *The Journal of Economic History* 58.2, pp. 345–374.

Haines, Michael R et al. (2010). “Historical, demographic, economic, and social data: the United States, 1790–2002”. In: *Ann Arbor, MI: Inter-university Consortium for Political and Social Research.*

Jaffe, Adam B, Mary Kaltenberg, Margie E Lachman, et al. (2021). *Invention and the Life Course: Age Differences in Patenting.* Tech. rep. Brandeis University.

Jones, Benjamin F (2009). “The burden of knowledge and the “death of the renaissance man”: Is innovation getting harder?” In: *The Review of Economic Studies* 76.1, pp. 283–317.

Kaltenberg, Mary, Adam B Jaffe, and Margie Lachman (2021). *The age of invention: Matching inventor ages to patents based on web-scraped sources.* Tech. rep. National Bureau of Economic Research.

Kelly, Bryan et al. (2021). “Measuring technological innovation over the long run”. In: *American Economic Review: Insights* 3.3, pp. 303–20.

Kogan, Leonid et al. (2017). “Technological innovation, resource allocation, and growth”. In: *The Quarterly Journal of Economics* 132.2, pp. 665–712.

Lamoreaux, Naomi R and Kenneth L Sokoloff (2005). *The decline of the independent inventor: a schumpterian story?*

Landes, William M and Lewis C Solmon (1972). “Compulsory schooling legislation: An economic analysis of law and social change in the nineteenth century”. In: *The journal of economic history*, pp. 54–91.

Lingwall, Jeff (2014). “An economic history of compulsory attendance and child labor laws in the United States, 1810-1926”. In: *Pittsburgh: Tepper School of Business. http://tepper.*

cmu. edu/~/media/files/tepper/extranet/academic% 20programs/phd/dissertations/lingwall_thesis_2014 pdf.

Lleras-Muney, Adriana (2002). “Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939”. In: *The Journal of Law and Economics* 45.2, pp. 401–435.

McAfee, Ward M (1998). *Religion, race, and reconstruction: The public school in the politics of the 1870s*. SUNY Press.

Moretti, Enrico (2021). “The effect of high-tech clusters on the productivity of top inventors”. In: *American Economic Review* 111.10, pp. 3328–75.

Petralia, Sergio, Pierre-Alexandre Balland, and David L Rigby (2016). “Unveiling the geography of historical patents in the United States from 1836 to 1975”. In: *Scientific data* 3.1, pp. 1–14.

Stephens, Melvin and Dou-Yan Yang (2014). “Compulsory education and the benefits of schooling”. In: *American Economic Review* 104.6, pp. 1777–92.

Toivanen, Otto and Lotta Väänänen (2016). “Education and Invention”. In: *Review of Economics and Statistics* 98.2, pp. 382–396.

A Appendix A

Appendix A provides details on the matching procedure between the 1940 Full Count Census and PATSTAT. Table A1 displays all name permutations for the name string: "GARRISON GEORGE HENRY."

Table A1: Permutations of "GARRISON GEORGE HENRY"

First Name	Last Name	Middle Initial
GARRISON	GEORGE	H
GARRISON	HENRY	G
GEORGE	HENRY	G
GEORGE	GARRISON	H
HENRY	GEORGE	G
HENRY	GARRISON	G

Notes: This table presents all permutations of the PATSTAT name string "GARRISON GEORGE HENRY", an inventor listed on US998585.

For the first group of Census names that have no duplicate first and last name in the Census we perform an exact match between the cleaned first and last names in the Census and patent records and retain the match as long as the middle initial does not contradict. The middle initial contradicts if the middle initials are both non-missing between the Census and PATSTAT and the middle initials do not match. We also drop matches where a first/last name permutation in the patent record has at least two different non-missing middle initials. The only exception to this rule is if the first name, last name, and middle initial all exactly match, then the match is retained.

Table A2 provides examples of hypothetical exact matches between first and last names in PATSTAT and the Census and whether the matches are kept based on the matching criteria outlined in the previous paragraph. In these hypothetical matches, the Census individuals have unique first and last names in the Census and all name strings in the patent record that share the listed first and last names are shown.[18] As an example of how middle initials

[18] For example, "Tyrie A Diggs", "Tyrie B Diggs", and "Tyrie Diggs" would be the only name strings that have a first name of "Tyrie", and a last name of "Diggs" in the patent record. Similarly, "Tyrie A Diggs"

can contradict, consider the match between "John L Doe" in the patent record and "John K Doe" in the Census. Since the middle initials of "K" and "L" contradict, then this match is not retained as shown in row 1 of Table A2. For an example of both criteria in action, consider the name string with first name "Tyrie" and last name "Diggs," shown in rows two through four of Table A2. His name appears three times in the patent record, twice with non-missing middle initials of "A" and "B" and once without a middle initial. There is only one "Tyrie Diggs" in the Census, and he has a middle initial of "A". In row two, "Tyrie A Diggs" from PATSTAT matches exactly to "Tyrie A Diggs" from the Census record, so the match is kept. In row three, the middle initial contradicts, so the match is not kept. In row four, the middle initial is missing in the patent record, so we do not know whether "Tyrie Diggs" should be associated with "Tyrie A Diggs", "Tyrie B Diggs", or another "Tyrie Diggs" with a different middle initial. As such, in row four we do not retain the match between "Tyrie Diggs" in the patent record and "Tyrie A Diggs" from the Census. In rows five through six, we see "Robert Adams" appears twice in the patent record with middle initials of "G" and "H." Since "Robert Adams" has at least two different non-missing middle initials in the patent record, the matches are discarded. This helps prevent multiple different persons in the patent record matching to the same Census individual. In rows seven and eight of Table A2, the first/last name permutation "Tommy Atkins" only appears with one non-missing middle initial in the patent record, "R." Since the middle initials do not contradict and there is only one non-missing middle initial for the patent name "Tommy Atkins", the matches in rows seven and eight are kept. While "Tommy Atkins" in the patent record could be different from "Tommy R Atkins" in the patent record, the likelihood of a false positive match should be significantly lower than when a first/last name permutation has multiple non-missing middle initials in the patent record.

The second group of Census individuals are those who have a non-unique first and last name in the Census, but all individuals with the same first and last name have non-missing middle initials. For these individuals we simply perform an exact match using their first name, last name, and middle initial.

is the only Census individual with first name "Tyrie" and last name "Diggs." In the Census, this is by construction as we only consider Census individuals with unique first and last names.

Table A2: Hypothetical Matches for Census Group #1

First Name	Last Name	MI Patent	MI Census	Keep the Match
John	Doe	L	K	No
Tyrie	Diggs	A	A	Yes
Tyrie	Diggs	B	A	No
Tyrie	Diggs		A	No
Robert	Adams	G		No
Robert	Adams	H		No
Tommy	Atkins		R	Yes
Tommy	Atkins	R	R	Yes

Notes: This table provides examples of hypothetical exact matches between first and last names in PATSTAT and the Census and whether the matches are kept based on our matching criteria. None of these names should be taken to reflect actual names of inventors in PATSTAT or names of individuals in the 1940 Full Count Census. In these hypothetical matches, the Census individuals have unique first and last names in the Census and all name strings in the patent record that share the listed first and last names are shown.

Table A3 displays summary statistics on the match, when splitting out individuals into three mutually exclusive groups: those who have no invalid matches (group 1), those who have all invalid matches (group 2), and those who have some invalid and some valid matches (group 3). The validity of the match is based on whether the individual's implied age at filing the patent is between 18-70. Although they only comprise 14% of individuals that match, persons with some valid matches and some invalid matches account for about 40% of all matches.

Table A3: Match Summary Statistics

	$\overline{\text{Invalid}} = 0$		$\overline{\text{Invalid}} = 1$		$0 < \overline{\text{Invalid}} < 1$	
	(1)		(2)		(3)	
	Mean	N	Mean	N	Mean	N
Age at Filing	43	250,315	39	50,095	54	212,472
Invalid Match	0	250,315	1	50,095	.47	212,472
$\overline{\text{Invalid Match}}$	0	68,469	1	19,221	.46	14,748
Matches	3.7	68,469	2.6	19,221	14	14,748

Notes: This table presents summary statistics on the implied age of inventor × patent observations that match to the 1940 Census, as well as summary statistics at the individual level. Statistics are grouped by whether the individual had no invalid matches (group 1), all invalid matches (group 2), or some invalid and some valid matches (group 3). "Age at Filing" is defined as the year of application for the patent minus the individual's birth year. "Invalid Match" is defined as a match where the "Age at Filing" is either less than eighteen or greater than seventy. "$\overline{\text{Invalid Match}}$" is defined as $\frac{1}{J}\sum_{j=1}^{J}\{\frac{1}{N}\sum_{n=1}^{N}\mathbb{1}\{\text{Invalid Match}_n\}\}_j$ where j indexes each individual who has at least one match with the patent record and n indexes that individual's patents. "Matches" is the number of matches an individual has with the patent record.

Appendix B

Table B1: The Effect of Compulsory Schooling on Lifetime Patenting (2SLS)

	Patents			
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Years of School	0.007***	0.007***	0.013	0.022
	(0.000)	(0.000)	(0.010)	(0.014)
Minority Race	0.005***	0.005***	0.022	0.046
	(0.001)	(0.001)	(0.029)	(0.038)
Patents per capita		-0.001		-0.002
		(0.002)		(0.002)
Urban share		-0.009**		-0.014**
		(0.004)		(0.006)
Mfg. Output per capita		0.005*		0.003
		(0.003)		(0.004)
$\overline{\text{Y}}$	0.039	0.039	0.039	0.039
KP *F*-Stat			17.07	9.64
Observations	6,687,254	6,687,254	6,687,254	6,687,254

Notes: This table presents results from estimating Equation 3 using 2SLS. "Patents" is a count of the lifetime number of patents granted to the individual. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth × year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).

Table B2: The Effect of Compulsory Schooling on Breakthrough Innovation (2SLS)

	Breakthrough Innovations (Top 20%)			
	(1)	(2)	(3)	(4)
	OLS	OLS	2SLS	2SLS
Years of School	0.001***	0.001***	0.002	0.005
	(0.000)	(0.000)	(0.003)	(0.003)
Minority Race	0.002***	0.002***	0.004	0.011
	(0.000)	(0.000)	(0.007)	(0.009)
Patents per capita		-0.000		-0.000
		(0.000)		(0.000)
Urban share		-0.002**		-0.003**
		(0.001)		(0.001)
Mfg. Output per capita		-0.000		-0.001
		(0.001)		(0.001)
$\overline{\text{Y}}$	0.007	0.007	0.007	0.007
KP F-Stat			17.07	9.64
Observations	6,687,254	6,687,254	6,687,254	6,687,254

Notes: This table presents results from estimating Equation 3 using 2SLS. "Breakthrough Innovations (Top 20%)" is a count of the lifetime number of breakthrough patents granted to the individual, which are patents which belong to the top 80 percentile of reverse-forward similarity in a five year window, as measured in Kelly et al. 2021. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth × year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).

Table B3: First Stage

	Years of School	
	(1)	(2)
	OLS	OLS
$\mathbb{1}\{\text{CS} = 6\}$	0.003	0.007
	(0.017)	(0.015)
$\mathbb{1}\{\text{CS} = 7\}$	0.052***	0.060***
	(0.019)	(0.018)
$\mathbb{1}\{\text{CS} = 8\}$	0.090***	0.083***
	(0.022)	(0.021)
$\mathbb{1}\{\text{CS} \geq 9\}$	0.181***	0.140***
	(0.027)	(0.028)
Minority Race	-2.773***	-2.773***
	(0.036)	(0.036)
Patents per capita		0.070***
		(0.020)
Urban share		0.285***
		(0.057)
Mfg. Output per capita		0.110***
		(0.040)
$\overline{\text{Y}}$	8.94	8.94
KP F-Stat	17.07	9.64
Observations	6,687,254	6,687,254

Notes: This table presents results from estimating Equation 2 using OLS. "Years of School" is a count of the number of years of formal schooling the individual achieved. "Minority Race" is a dummy for all races except "White." Patents per capita (1,000 persons), urban share, and manufacturing output per capita are all measured in the individual's state of birth in the year the individual turns 18. These variables are standardized to have mean zero and standard deviation of one by year. Standard errors are clustered at the state of birth × year of birth level and shown in parentheses. *(p<0.1), **(p<0.05), ***(p<0.01).