The Long-Run Effects of Parental Wealth Shocks on Children

Cache Ellsworth Ian Fillmore Adrian Haws Joseph Price

July 3, 2025

Abstract

Despite its central importance in economics and public policy, there is limited evidence of the long-run causal effects of parental wealth. We leverage a natural experiment to provide new evidence on this question. Beginning with the land run of 1889, homesteaders in Oklahoma Territory raced to claim plots of land, unaware of oil fields hidden beneath their feet. Over the next few decades, some homesteaders received a stream of royalties from oil production while less fortunate neighbors did not. We construct an intergenerational panel dataset that links land and oil records to U.S. census records and tracks homesteaders and their children across multiple decades. Our identification strategy leverages highly local variation in the boundaries of oil fields to compare individuals within a six-by-six mile survey township, separating the individual wealth shock from the community effects of oil discoveries. We find that homesteaders with oil on their land retired at higher rates and moved to towns and cities within Oklahoma. Their children who were under age 18 at the time of oil discoveries received an additional 1.1 years of education (a 10% increase) and were 17 percentage points more likely to graduate from high school (a 37% increase). Sons of homesteaders with oil wealth were more likely to be in the labor force, and worked in occupations associated with 14% higher incomes. Daughters were 11 percentage points less likely to have been married by the 1940 census, likely indicating delayed marriage. We also find evidence suggesting that daughters married more educated spouses. Our study is the first to identify the distinct effects of parental wealth for daughters, and may reconcile previous evidence on the effects of wealth on children's education.

JEL Codes: J24, J62, N32, O13

This research is supported by the Weidenbaum Center at Washington University in St. Louis. We thank Thomas Barden, Zach Flynn, Jiangheng Lin, and Bryson Mumford for excellent research assistance. We are especially grateful to Micah Villarreal for helpful discussions on data sources. Authors may be contacted by email at crellsworth@wisc.edu, ianfillmore@gmail.com, aah226@cornell.edu, and joseph_price@byu.edu.

1 Introduction

Economic theorists have long explored the question of whether unequal advantages persist across generations.[1] This question is also at the center of historical and modern policy debates.[2] Despite its theoretical and practical importance, there is relatively little causal evidence on the effects of parental wealth on children's long-run outcomes. A large literature establishes empirical correlations between parents' and children's wealth and income, but these studies are generally not designed to isolate parental wealth from its many correlates.[3] Although several studies address causal effects (Lovenheim, 2011; Lovenheim and Reynolds, 2013; Cesarini et al., 2017; Bleakley and Ferrie, 2016; Ager et al., 2021; Bulman et al., 2021; Daysal et al., 2023), there is conflicting evidence on key outcomes such as education as well as limited or no evidence on marriage, migration, and income.[4] Additionally, no studies analyze heterogeneous effects by gender, and existing historical studies omit daughters entirely. Studying effects for daughters is a substantial gap in the existing literature, especially because of historical gender gaps in earnings, occupations, and wealth (Goldin, 1990; Ruel and Hauser, 2013).

In this paper, we leverage a natural experiment in early 20th-century Oklahoma to provide new evidence on the causal effects of parental wealth on sons' and daughters' labor supply, income, education, migration, marriage, and wealth. During the years following Oklahoma's founding in 1907, oil discoveries began dotting the state. Petroleum reserves provided instant access to wealth for a large number of private landowners, including many farmers who had claimed 160-acre plots through the Homestead Act. "Wildcatters" drilled exploratory wells, offering substantial royalties to farmers in return for leasing the land. The median size of these "wealth shocks" was about $14,000 in 1940, which was 28 times the yearly earnings of a manual

[1] Mogstad (2017) and Becker et al. (2018) are recent examples.

[2] One example is a speech on economic mobility by President Obama (Obama, 2013).

[3] Recent work finds that between one-quarter and three-quarters of the variation in economic status for children born between 1840 and 1980 can be explained by their parents' economic status ((Ward, 2023; Davis and Mazumder, 2024; Jácome et al., 2025)). Also see Black and Devereux (2011), Stuhler and Biagi (2018), and Abramitzky et al. (2025) for review articles on intergenerational persistence.

[4] A larger literature studies within-generation effects of wealth from money and land lotteries (Bleakley and Ferrie, 2016; Cesarini et al., 2016, 2017; Lindqvist et al., 2020; Kim and Koh, 2021; Golosov et al., 2024), house price booms (Lovenheim and Mumford, 2013; Aladangady, 2017), fracking booms (Bellon et al., 2021; Cookson et al., 2022), and inheritances (Elinder et al., 2018; Nekoei and Seim, 2023) on consumption, labor supply, debt repayment, entrepreneurship, fertility, mental and physical health, and mortality.

laborer.[5] Two specific aspects of these wealth shocks allow us to identify causality. First, geographic variation in oil deposits allows us to compare neighbors with or without a wealth shock, removing concerns of unobserved regional characteristics correlated with the presence of oil. This comparison also nets out the local effects of a resource boom.[6] Second, homesteaders did not have the technology to predict whether a particular plot of land would have oil.[7] We also show that homesteaders' characteristics are balanced across treatment status.

The timing of these wealth shocks is ideal for studying a broad range of long-run outcomes for both sons and daughters. Many children of homesteaders were under age 18 at the time of oil discovery, allowing us to study effects on human capital and marriage outcomes. Additionally, most of the children in our sample were ages 18-40 in the 1940, allowing us to directly observe their education, migration, marriage, fertility, income, labor supply, occupational choice, and wealth using the rich information in the 1940 U.S. Census.

One challenge in constructing this dataset is that homesteaders' land titles contain no biographical information other than names. We develop a strategy for sparse record linkage, leveraging information on homesteaders' neighbors to increase confidence in a potential link. This strategy allows us to successfully link over 60% of homesteaders to the 1900 or 1910 U.S. census, which is much higher than linkage rate for settings that have similar sparse data such as inventor patents which achieve a match rate of 10% (Sarada et al., 2019). Balancing tests find little to no evidence of differences across neighboring homesteaders with or without potential access to oil. We then use linked census data from the Census Tree Project (Price et al., 2021; Buckles et al., 2024) to find homesteaders' children and link them to the 1940 Census.

[5] Wealth could be received as a perpetual stream of royalties, or by selling mineral rights associated with the land. We calculate the size of the wealth shock as the net present value of production royalties. See Section A.3 of the Online Appendix for more details.

[6] Several studies have examined the local economic effects of discovering and extracting natural resources, finding permanent population increases (Brown and Cuberes, 2020), higher employment and wages (Black et al., 2005a; Feyrer et al., 2017; Bütikofer et al., 2025) and lower human capital investment (Black et al., 2005b; Mousavi and Clark, 2021). Studies additionally find negative effects on local governance (Cust and Poelhekke, 2015) and varying effects on house prices (Muehlenbachs et al., 2015).

[7] Owen (1975) notes that "occasional sensational discoveries were accompanied by an inordinate number of failures on prospects that looked as favorable as the richest productive spots." To solve this problem, geophysical exploration techniques were developed around the 1920s (Fulton and Stringer, 2009), long after the homesteaders in our sample acquired land. Yet "despite the amazing evolution of technology and exploration techniques, the only way of being sure that a reservoir exists is to drill an exploratory well" (Department of Transportation, 2025).

The Census Tree relies on crowd-sourced genealogy data to provide a large number of links for daughters, who were previously omitted in related studies. We also observe short-run outcomes for parents in the 1930 Census.

We first examine the effects of wealth on parents' short-run outcomes using data from the 1930 US Census. We use an intention-to-treat approach since our random variation on access to oil wealth is based on using homesteading decisions that were made prior to the oil discoveries (with some of the homesteaders in our sample migrating away before oil was discovered). We find that homesteaders with oil on their land were 11.6 percentage points less likely to be in the labor force (relative to 74.8% for the control group), suggesting an increased likelihood of retirement. This group was 13.3 percentage points more likely to be living in a town or city (a 52% increase) and about 5.1 percentage points less likely to be living in Oklahoma (though not statistically significant). These effects are measured in 1930, so parents' migration occurred before the Dust Bowl.

Wealth has been shown to increase fertility (Lovenheim and Mumford, 2013; Bleakley and Ferrie, 2016), leading to a question of whether the effects of parental wealth are diluted among an increased number of children. However, the age distribution of homesteaders in our sample makes this an unlikely explanation.[8] Additionally, it is unlikely that outcomes such as public education are zero-sum among children in the same family.

Parental wealth shocks during childhood had dramatic effects on children's education. Sons and daughters gained 1.1 additional years of education, relative to a baseline mean of approximately 10 years. The effect of wealth on high school graduation is especially pronounced for sons, who were 20.1 percentage points more likely to complete high school. Daughters were 11 percentage points more likely to do so. Because daughters in the control group received more education, these effects worked toward reducing the gender gap in educational attainment.

The evidence on children's labor market and wealth outcomes is mixed. Sons of homesteaders with oil wealth were 4.9 percentage points more likely to be in the labor force and earned 14% more occupation-based income, but we find no effect on these outcomes for daughters. Although

[8] As shown in Table 1, the average age of homesteaders in 1900 was approximately 40, and the earliest oil discovery in our sample was in 1911. Figure 6 further shows that nearly all of the children in our sample were born before oil was discovered in their community.

we find no statistically significant effect on increased earnings for sons or daughters, estimates are large and noisy because the subsample becomes very small. We indirectly observe effects on children's wealth through home ownership. There is no effect on home ownership, but we again find large and imprecise estimates that suggest increased home values among the subsample.

The effects of parental wealth on children's marriage outcomes are also striking. We find that daughters of treated homesteaders are 11 percentage points less likely to have ever been married by 1940, while there is no detectable effect on marriage outcomes for sons. The children in our sample were ages 18-40 in the 1940 census, so this effect suggests that daughters delayed marriage. We also find suggestive evidence that daughters married spouses with more education.

Our analysis provides empirical evidence that wealth can have dramatic effects on human capital for both sons and daughters, supporting prior theory (Mogstad, 2017). The especially large effect of parental wealth on sons' probability of finishing high school could be driven by less dependence on child farm labor as wealth led parents to retire early and exit farming. Parental migration could have also indirectly affected children's education through increased access to schooling. Human capital likely explains the income effect we observe for sons.

Relation to Prior Literature — Several existing studies identify intergenerational effects of wealth. Bleakley and Ferrie (2016) observe the universe of winners and losers from a land lottery in Antebellum Georgia, finding no effect on the school attendance, literacy, proxied income, or wealth of sons. Ager et al. (2021) study the effects of emancipation on the sons of former slaveholders, finding that a negative wealth shock had limited effects on sons' proxied wealth. A series of studies use housing price variation to document positive effects of parental wealth on children's wealth (Daysal et al., 2023), college attendance (Lovenheim, 2011), and college choice (Lovenheim and Reynolds, 2013). Wealth from lotteries also increases children's college attendance (Bulman et al., 2021), but does not affect school performance (Cesarini et al., 2016). A related literature studies the effects of parental income on children's schooling and health, mostly in the short run.[9]

[9] Many of these studies focus on tax credits (Chetty et al., 2011) and cash transfers (Aizer et al., 2024; Baird et al., 2014; Pega et al., 2022). Bastian and Michelmore (2018) study long-run effects of the Earned Income Tax Credit, finding effects on children's high school and college completion, employment and earnings.

Our work is concurrent with Villarreal (2025), who studies the intergenerational effects of oil discoveries among Creek Freedmen in former Indian Territory. This population had descended from slavery and faced barriers to economic mobility caused by segregation and racial violence.[10] Despite differences in the population that received wealth shocks, Villarreal (2025) finds broadly similar effects on migration and children's education. Our paper differs in two additional respects. First, our identification strategy is designed to isolate the effects of individual wealth from the local effects of resource booms. Second, we study heterogeneous effects by gender.

Our study contributes to the literature by providing clarifying evidence of the effects of wealth on educational attainment. Although studies in modern contexts find that parental wealth increases children's college attendance (Lovenheim, 2011; Bulman et al., 2021), Bleakley and Ferrie (2016) find no effects on school attendance or literacy. By contrast, we find large effects of wealth on high school graduation and educational attainment. These different effects may be driven by the size and type of the wealth shock, as well as the historical context. Winners of the Georgia land lottery discussed in Bleakley and Ferrie (2016) may have still relied on farming because the size of the wealth shock was smaller and winners could likely only benefit through selling their land or investing in agricultural production.[11] Despite the increase in wealth, it is not clear that lottery winners would send their children to school more if they relied on them to help with farming. A passive stream of oil royalties, on the other hand, led homesteaders to exit the labor force and may have reduced the demand for child labor. Our findings suggest that wealth can affect demand for education, even in a context with public education and compulsory schooling (Davison, 1950).

Our study is the first to identify the distinct effects of parental wealth for daughters. We find that parental wealth shocks delay marriage for daughters, consistent with the observational findings of Axinn and Thornton (1992). We also find that parental wealth causes daughters to marry husbands who are more educated. In contrast, we find no effect on marriage outcomes for sons. This finding provides context to a large literature on assortative mating and suggests

[10]The 1921 Tulsa Race Massacre razed a prominent Black community living within the boundaries of the Muscogee (Creek) Nation Reservation, with long-run economic consequences (Albright et al., 2021).

[11]The practical value of land received in the Georgia land lottery, 5 years of wages for a manual laborer, was less than one fifth of the median value of oil royalties in our sample.

that the channels by which parental wealth is transmitted to children may vary by gender.[12]

We also contribute to the literature on within-generation wealth effects. Kim and Koh (2021) find higher consumption, but no effect on labor supply from lottery wins. Bleakley and Ferrie (2016) find higher fertility. Cesarini et al. (2016) find no effect on healthcare utilization or mortality. Cesarini et al. (2017) find modest but persistent declines in income. Golosov et al. (2024) also find reduced labor supply and increased consumption, as well as migration to better neighborhoods within the same state. We corroborate evidence on reduced labor supply and increased migration.

Finally, our paper contributes to methods for linking land records which contain only sparse biographical information. A growing literature studies the economic history of homesteading and land allotments in the U.S.[13] Hauck (2023a) and Villarreal (2025) have linked land records to individual-level census data, but the full potential of these data sources has yet to be realized. We provide an automated and reproducible linking strategy that leverages information about neighbors to obtain high match rates despite sparse data.

2 Historical Background

The historical events in Oklahoma provide a natural experiment, allowing us to estimate causal effects of parental wealth. We provide some of the first causal evidence on the effects of parental wealth on education, labor market, and marriage outcomes, while also supporting previous findings from time periods and places very different from 20th-century Oklahoma.

2.1 Homesteading in Oklahoma

The Homestead Act of 1862 played a key role in facilitating westward population expansion. Settlers willing to cultivate the land received 160 acres from the federal government for only a small filing fee. Most of present-day Oklahoma had been set aside for the forced resettlement

[12] Studies on assortative mating have documented strong correlations between the education, wealth, and family background of spouses (Kalmijn, 1998; Charles et al., 2013; Eika et al., 2019).

[13] These articles include Allen and Leonard (2020), Leonard et al. (2020), Allen and Leonard (2021), Leonard and Parker (2021), Mattheis and Raz (2021), Leonard and Kogelmann (2022), Hauck (2023a), Hauck (2023b), Hauck and Woutersen (2023), Dippel et al. (2024), and Villarreal (2025).

of Native Americans, including the five tribal nations removed during the "Trail of Tears", but would-be homesteaders had an increasing appetite for Indian Territory by the 1880s. Congress authorized homesteading on the so-called Unassigned Lands which had not been assigned to tribal nations, and the first "land run" was held on April 22, 1889. On May 2, 1890, Oklahoma Territory was created from the western half of present-day Oklahoma (Indian Territory remained in the eastern half). This was followed by "assimilation" policies such as the Dawes Act that opened large amounts of tribal lands for homesteading.[14] The newly-created Oklahoma Territory was one of the last states or territories to be opened for homesteading, and it represented nearly one in six homestead claims between 1889 and Oklahoma statehood in 1907 (Wilm et al., 2025). Figure 2 shows the timeline and geography of land runs and land lotteries in Oklahoma Territory.

Homesteaders were required to remain on the land for five years before they could receive ownership via a land patent. Our sample consists of homesteaders who satisfied the requirements and received a patent, rather than abandon their homestead in the first few years. We do not study effects of oil discoveries on Native Americans in this paper because land allotments were distributed in a much less random way and Indian Territory contained a patchwork of individual allotments, communal tribal lands, and white-owned homesteads.[15]

We study the effects of wealth during a dynamic period. Children in our sample were exposed to the high school movement, which dramatically increased access to secondary schooling (Goldin, 1998). We also observe children's outcomes in the wake of the Dust Bowl and the Great Depression. The Dust Bowl is a salient part of Oklahoma history, providing the background for the "Okie" migrants in *The Grapes of Wrath* (Steinbeck, 1939).[16] Our empirical strategy is equipped to account for these historical events. We compare individuals within townships, which are small enough to account for local access to schools, local economic characteristics, and geographic variation in exposure to natural disasters. Additionally, we include a full set of birth cohort fixed effects, which account for the trend in high school construction. Although they do

[14]The Dawes Act of 1887 divided tribal land into individual allotments of different sizes based on age and marital status. Remaining plots were then opened to homesteading.

[15]Anecdotal evidence suggests that many Native Americans became extremely wealthy from oil, but exploitation from the "Osage murders" partially resulted from the system of hereditary headrights to mineral wealth (Grann, 2018).

[16]Economic analysis has provided additional evidence of the effects of the Dust Bowl on land (Hornbeck, 2012) and populations (Arthi, 2018; Hornbeck, 2023).

not pose a threat to our analysis, these events provide important context for our study.

2.2 Oklahoma oil discoveries

During the first half of the 20th century, U.S. oil production grew more than thirty-fold in response to the rise of the automobile (U.S. Energy Information Administration, 2025).[17] Much of this oil boom was driven by newly discovered oil deposits in Oklahoma and neighboring states.[18] Figure 1 shows that oil deposits in Oklahoma were widespread.[19] Through sheer luck, some ordinary farmers benefited from oil discoveries on their lands. Yet because these oil deposits were extremely local, not all of their neighbors were as fortunate.

The story of Frank Wheeler illustrates how this unexpected wealth changed the lives of Oklahoma landowners (Franks, 1980, pp. 68–71).

> By the time Oklahoma entered the Union in 1907, Frank Wheeler had already purchased a farm about twelve miles east of the community of Cushing... There, he shared a small log house with his wife, eight daughters, and one son. The rough terrain was described as "sixty acres of plow land and one hundred acres of rocks and scrub." It was difficult for Wheeler to wrest a living from the barren earth, and he spent much of the time seeking employment as a mason to supplement his income. Even so, the family remained constantly on the brink of poverty....
>
> ...Pennsylvanian Tom Slick arrived in Cushing during the winter of 1911.... Slick was known for drilling wildcat wells, and many potential investors were skeptical of his idea of oil in the area, pointing out that the nearest producing well was more than twenty-six miles away....

[17]The U.S. produced 164 million barrels of crude oil in 1900, which grew to 1,974 million in 1950. By 1950, 42 years after the release of the Model T, there was nearly one passenger car per U.S. household (Mitchell, 1993).

[18]Owen (1975) remarks: "In 1920, the Mid-Continent's production of 250,111,000 barrels was more than 56 percent of the nation's total... A phenomenal series of discoveries supplied most of the oil—Electra, north Texas, in 1911; Cushing, Oklahoma, in 1912; Healdton, Oklahoma, in 1913; Augusta, Kansas, in 1914; El Dorado, Kansas, in 1915; Garber, Oklahoma, in 1916; Ranger, Texas, in 1917; Breckenridge and Burkburnett townsites, Texas, in 1918; Northwest Burkburnett, and Hewitt, Oklahoma, in 1919; Burbank, Oklahoma, in 1920—but many smaller pools in the same regions were substantial contributors."

[19]Many of these reservoirs contained both petroleum and natural gas. We refer to all wells as "oil wells" for simplicity because most wells produced oil and oil was far more lucrative at the time.

> Although somewhat hesitant, Wheeler, after being informed of Slick's reason for being in the area, agreed to lease the land.
>
> . . . on March 12, 1912, [the well] blew in "with a roar of gas-driven oil that literally flooded the surrounding earth." The tremendous pressure of the natural gas in the well forced the crude to the surface at a rate of 400 barrels a day. The great Cushing Pool had been discovered. Within a month, Wheeler was collecting $125 daily [over $4,000 in 2025 dollars] in royalties from the barren 160 acres that had produced such a meager existence for his family only days previously.

Oil prospectors such as Tom Slick were motivated by huge potential profits of finding new oil deposits, but they also took a risk. Wells were very costly, and exploration was a very uncertain process with many wells coming up dry.[20] Because wildcatters or oil companies paid the drilling costs and shouldered the risk, the wealth was truly passive for landowners. As the Cushing Independent noted in 1912, "land owners have everything to gain and no risk to themselves in making leases" (Cushing Independent, 1912). A standard lease agreements awarded a royalty of 12.5 percent to the land owner.

3 Data

3.1 Oklahoma Geological Survey

We identify homesteaders who likely had access to oil wealth by overlaying the geography of oil wells with the locations of federal land patents. Figure 1 shows the locations of 119,519 oil wells completed by 1940, from the Oklahoma Geological Survey (Boyd, 2025). These data contain the timing of completion and the exact coordinates of each well. We identify wells that were completed by 1940 within each quarter-section (160 acres) in the Public Land Survey System. A key aspect of these discoveries is that they are extremely localized. Because we are unable to distinguish between dry or productive wells, we isolate the comparison to homesteaders who lived on a plot of land with 4 or more wells to those who had no wells. This comparison removes

[20] The Cushing Independent noted that "it costs from $8,000 to $10,000 to put down a single hole" (Cushing Independent, 1912).

much of the uncertainty around access to wealth. A quarter-section with fewer than 4 wells likely indicates exploratory well that produced little oil or gas.

3.2 Federal Land Patents

Data on homesteaders is drawn from the Bureau of Land Management (BLM) Federal Land Patents (U.S. Department of the Interior, 2022), which identify the name, precise location of land ownership, and patent date for over 72,000 land patents (equivalently, deeds or titles) transferred from the federal government to individual homesteaders in Oklahoma Territory by 1910. Prior to statehood in 1907, Oklahoma Territory was opened to rapid settlement starting in 1889 (see Figure 2). We focus on homesteaders with a land patent between 1889 and 1910 in former Oklahoma Territory; see the map in Figure 3. As shown in Figure 4, these homesteaders obtained land ownership before oil discoveries in those areas. We focus on successful homestead entries to reduce attenuation bias due to out-migration of abandoned homesteads and to eliminate any selection bias from strategic land purchases.[21] We use the Public Land Survey System (PLSS) grid to assign homesteaders to the 6x6 mile township where they received a land patent. Survey townships could accommodate 136 homesteads, each with a 160-acre quarter-section.[22] This allows us to compare homesteaders with or without access to oil within a very small geography, about 1/25 of the average size of counties in present-day Oklahoma.

3.3 Census linking

Figure 5 illustrates our sample construction process. We first link land patents to individual census records in 1900 and 1910. This step is critical for identifying family relationships, but it presents a difficult record linking challenge because of the sparse biographical information available in the land patents. We develop a strategy that leverages information on neighbors in

[21] After filing for an initial homestead, there were three possible outcomes—"prove" (obtain ownership after fulfilling the requirements), abandon, or commute the homestead. To successfully prove a homestead, individuals applied for a patent after at least five years of "settlement and cultivation" (U.S. National Archives and Records Administration, 1862). Up to 45 percent of initial homestead claims were abandoned and returned to the federal government (Hauck, 2023a). Some individuals instead chose to commute their homesteads by purchasing them before the five years were completed. More details are included in Appendix A.

[22] Although a township contains 36 sections (144 quarter-sections), two sections were initially set aside in each township for financing or constructing schools.

both the land patents and the census, resulting in 59% of land patents linked to a census record. Our neighbor strategy provides a huge advantage over other sparse data linking approaches which try to link datasets with only name, town, and year to census records. For example, efforts to link inventor patents to census records often achieve match rates of only 10%, such as in Sarada et al. (2019). See Appendix B for more details on the linking procedure.

The Census Tree (Buckles et al., 2024) provides the core data allowing us to link homesteaders to their children's outcomes in adulthood.[23] This dataset is the largest set of linkages among U.S. Census records from 1850 to 1940, with over 700 million individual linkages including both men and women.[24] We use the Census Tree to link homesteaders from the 1900 or 1910 U.S. Census to other census years and observe their household family relationships, where we then collect each homesteader's children in each census year.[25] For example, if we observe a homesteader in the 1880, 1900, and 1910 censuses we gather all children living with the homesteader in each of those years. We then link each of these children forward to the 1940 Census to de-duplicate individuals and to observe their outcomes in adulthood. A similar process allows us to link homesteaders to their own outcomes in the 1930 Census.

3.4 Characteristics of the linked data

Table 1 shows that we find no evidence of systematic differences in the characteristics of homesteaders across treatment status. Within the same township, homesteaders with at least 4 oil wells on their land (by 1940) do not differ by socioeconomic status or demographics, across 15 characteristics. This underscores the historical evidence that homesteaders could not predict whether oil would eventually be found on their land.

We restrict the sample of children to those whose parent resided in a township that had oil by 1940. Additionally, we focus on children who were under age 18 when the median oil discovery occurred within the township. This restriction is important for two reasons. First, it minimizes the likelihood that children were directly affected by a wealth shock on their own

[23] See also Price et al. (2021).

[24] Census data is provided by IPUMS USA (Ruggles et al., 2024a,b).

[25] Several different linking approaches form the Census Tree links, and the most accurate links are found by multiple sources (Buckles et al., 2024). We ensure a high level of accuracy among these links by dropping links from only one source.

land.[26] Second, it allows us to study effects in young adulthood that are theoretically driven by childhood investments and decisions, including education, occupational choice, migration, and marriage. Finally, we restrict the sample of children to those who were at least 18 years old in 1940. Age distributions of children and parents are shown in Figure 6.

4 Empirical Strategy

We compare the children of neighboring homesteaders with and without oil. Since the homesteading decisions were made prior to oil discoveries, this comparison of neighbors allows us to assume that the individuals and the land they reside on are similar. We implement this comparison among neighbors by estimating a regression with township fixed effects. Focusing on variation within townships makes our identification strategy more credible because geography, climate, and economic characteristics are unlikely to vary greatly at such a local level. Specifically, our identifying assumption is that, within a township, the presence or absence of oil on a homesteader's land is uncorrelated with any unobserved characteristics of the homesteader. Although we cannot directly test this assumption, we do verify that homesteaders with oil are observably similar to their neighbors without oil (see Table 1). Additionally, we find that the presence of oil is unrelated to the likelihood of linking homesteaders to the census (see Table B1 in the Appendix). Taken together, this evidence supports our identifying assumption that the presence or absence of oil among early homesteaders was as good as randomly assigned.

Measurement error in the treatment variable is a challenge for our analysis. Some of the completed oil and gas wells in the data were actually exploratory wells that came up dry, but we do not observe this. If we were to assign homesteaders with any wells as "treated", our estimates would be severely attenuated because of false positives. To address this challenge, we compare homesteaders with multiple wells (4 or more) to those with no wells. Thus, we focus on the observations with less measurement error. A possible weakness of this approach is that our estimates may not include homesteaders with access to small amounts of wealth; however,

[26] Individual wealth shocks would not be a concern if they impacted the treated and control group equally and if individual and parental wealth shocks were not interactive. Neither of these is likely the case. Adult children could have received or purchased land near their parents, so their individual likelihood of oil would be correlated with their parents.

historical maps indicates that wells were located very close together, so 4 wells is not exceptional (see Figure A1 in the Online Appendix).Regulations by 1935 required no more than one well per 10 acres, meaning that a 160-acre homestead could have up to 16 wells. Prior to 1935, some homesteads had dozens of wells. Additionally, our definition of access to oil results in a geographic distribution that is similar to historical maps of oil field boundaries (see Figure A2 in the Online Appendix). We include observations with fewer than 4 wells in the sample, but include a separate dummy variable in our regressions for these observations with 1-3 wells. This ensures that our key comparison of interest is between the treated and control observations.

4.1 Estimating the effects of oil discovery

We begin by estimating the effect of oil discovery on the homesteaders themselves. We do this by linking each homesteader to the 1930 census and estimating the following regression specification

$$y_{jt} = \gamma_0 + \gamma_1 Treated_j + \gamma_2 Unsure_j + X_j \cdot \beta + a_t + \epsilon_{jt} \tag{1}$$

where y_{jt} is an outcome in 1930 for homesteader j living in township t, $Treated_j$ is a dummy for whether the homesteader likely had oil on their land by 1930 (as measured by having 4 or more wells), $Unsure_j$ is a dummy for whether the homesteader had uncertain access to oil by 1930 (as measured by having 1-3 wells), X_j includes homesteader covariates (gender and birth cohort dummies), and a_t is a set of township fixed effects.[27] The parameter γ_1 represents the effect of an oil discovery on outcome y for homesteaders in 1930.

Our main estimates for children come from the following regression specification

$$y_{ijt} = \delta_0 + \delta_1 Treated_j + \delta_2 Unsure_j + X_i \cdot \beta + a_t + \epsilon_{ijt} \tag{2}$$

where y_{ijt} is an outcome in 1940 for child i of homesteader j living in township t, $Treated_j$ is a dummy for whether homesteader j likely had oil on their land by 1940, $Unsure_j$ is a dummy for whether the homesteader had uncertain access to oil by 1940, X_i includes child covariates (gender and birth cohort dummies), and a_t is a set of township fixed effects. The parameter

[27]Homesteaders with oil after 1930 but not before are included in the "control" group.

δ_1 represents the effect of an oil discovery on outcome y for the children of homesteaders in 1940. By including township fixed effects, our estimator compares the children of homesteaders with oil to the children of their neighbors who did not have oil. We follow de Chaisemartin and Ramirez-Cuellar (2024) by clustering our standard errors at the township level.[28]

5 Effects of Wealth on Homesteaders

We first estimate the effects of oil wealth on the homesteaders in our sample (the parents) using data from the 1930 U.S. Census. Figure 6 shows that many homesteaders had surpassed the Social Security retirement age of 65 (Costa, 1998). Table 2 reports intent-to-treat estimates because some homesteaders may have migrated before oil was discovered on their land. As such, our estimates represent a lower bound of the effects of the wealth shock because of attenuation from homesteaders who weren't actually "treated" because they migrated prior to the oil discovery.

We find that access to oil wealth reduced the likelihood of being in the labor force by 11.6 percentage points, relative to a control group mean of 74.8% in the labor force. We estimate a 6.6 percentage point reduction in the likelihood of being a farm owner-operator or tenant farmer, which is only statistically significant at the 10% level. These effects likely indicate that wealth made it possible to retire, and many individuals retired from farming. Wealth also lowered the occupational income score (imputed income) of homesteaders by approximately \$223 (in 1950 values), which is 16% lower than the control group mean.[29]

We also estimate that homesteaders with oil wealth were 13.3 percentage points more likely to live in a town or city, a 52% increase relative to the control group. Our point estimates for staying in Oklahoma, indicate that homesteaders with oil wealth were 5.1% less likely to stay in Oklahoma, though this result is not statistically significant. Census day in 1930 occurred on April 1st, so all of our measures of migration were just prior to the start of Dust Bowl.

[28] Bell and Mccaffrey (2002) develop a small sample correction, which is often applied for fewer than 50 clusters (Cameron and Miller, 2015). In our case we have approximately 100 townships that contain both treated and untreated children, so this correction is not necessary.

[29] The occupational income score (provided by IPUMS USA as "occscore") is the median total income within each occupation among a sample from the 1950 Census (Ruggles et al., 2024a). The score is measured in hundreds of 1950 dollars.

6 Long-Run Effects of Wealth on Children

We now turn to the long-run effects of parental wealth on children's outcomes using the rich data available in the 1940 U.S. Census. Outcome categories include labor supply, income, education, migration, marriage, and home ownership. We provide intent-to-treat estimates for both the full sample of children and separately for sons and daughters. Our sample for this analysis includes 1,443 individuals, of which 201 experienced a likely oil wealth shock before the age of 18.

6.1 Labor supply and income

We first explore the effects of parental wealth on children's labor supply and income in adulthood, as reported in Table 3. Parental wealth increases the likelihood of being in the labor force by 4.9 percentage points for sons, with no statistically significant effect for daughters. Sons' increased labor force participation stands in contrast to the decreased labor force participation of their parents. We find no effect of parental wealth on the likelihood of being a farm owner-operator or tenant farmer.

The 1940 Census reports wage income but not business or farm income. We first provide estimates for an occupational income score, which is the imputed total income by occupation.[30] Parental wealth raised the income score of sons by approximately \$250 dollars (in 1950), which is 13.5% higher than the control group mean. This implies that sons held occupations with a higher typical income. There is no statistically significant effect of this outcome for daughters (which is less informative since labor force participation rates were lower for women during this time).

Measuring occupational income is possible for the full sample, but it does not capture individual-specific income differences that may vary by the wage rate and labor supply. To account for these, we also provide estimates for the log annual earnings among the subset of children who worked for wages.[31] Focusing on wage workers shrinks the sample considerably, partially driven by the prevalence of self-employed farmers. We do not find statistically sig-

[30] The occupational income score is described in footnote [29].

[31] We restrict the sample to children who had positive earnings and who were not unemployed, self-employed, or unpaid family workers at the time of the census. Wage income is reported for 1939, so we also require children to have worked at least 36 weeks in 1939.

nificant effects of parental wealth on log annual earnings for sons or daughters. However, the estimated coefficients are large and statistical power is limited due to the reduced sample size.

6.2 Education and migration

Table 4 demonstrates that parental wealth dramatically increased children's education. Sons and daughters each gained 1.1 additional years of education by 1940. These effects are strongly significant and amount to a 10-11% increase relative to the control group.

In 7, we test the robustness of our education results to the age cut-off that we use for children when assigning whether they were treated by the oil wealth shock experienced by their parents. If we expand our sample to include children who were in their twenties at the time of oil discovery, we find that children who were under 21 at the time of local oil discovery attain roughly one additional year of schooling. The effect becomes much smaller in magnitude when we focus on children whose family oil wealth shock occurred when they were age 26-29, after the typical age of college attendance. The null effect for this older group provides a nice placebo test for our main results.

We further explore education effects by focusing on the probability of completing high school. High school completion was a significant milestone which only 46% of the control group achieved. Sons of homesteaders with oil wealth were 20.1 percentage points more likely to complete high school (a 49% increase relative to the control group). For daughters, the increase in high school completion was 11 percentage points (a 21% increase), though the estimate for daughters is only statistically significant at the 10% level. Among our control group, daughters were 11 percentage points more likely to complete high school. As such, the suggestive evidence of differential results by gender indicate that these oil wealth shocks may have dramatically reduced the pre-existing gender gaps in high school completion.

We also examine the effects of parental wealth on children's migration, finding no statistical difference in their likelihood of urban residence or living in Oklahoma. However, the estimates of the likelihood of urban residence are in the same direction as the effect for their parents. Also, while the coefficient for moving to a larger city is imprecise, it suggests that the children in families that experience an oil wealth shock were 6-7 percentage points more likely to move to a large city (a 18% increase relative to the control group).

We consider two possible explanations for the effect of parental wealth on increasing children's education. The first is that parents who stop farming and exit the labor force may have a lower opportunity cost of sending their children to high school or college. Farming households required a great deal of labor, and many children worked informally on the farm during childhood and young adulthood. A passive stream of wealth could reduce the labor demands on both the parents and the children, freeing children to invest in their human capital.

The second explanation is that children move with their parents to towns and cities, which provide increased access to high schools and colleges. Parents with oil wealth are move likely to live in urban areas in 1930, at a time when many children in our sample were at the appropriate age for secondary and post-secondary schooling.

6.3 Marriage

Table 5 reports that parental wealth lowered daughters' likelihood of marriage and may have affected the characteristics of the spouses of daughters who did marry. Daughters of homesteaders with oil wealth are 11 percentage points less likely to ever have been married by 1940. Because many of the women in our sample were younger than age 30 in 1940, it is likely that this effect implies delayed marriage rather than lower likelihood of eventual marriage.

Among those children who were married and living with their spouse in 1940, We examine the characteristics of their spouse. We find weak evidence to suggest that children married more educated spouses. Among the combined sample of sons and daughters, parental wealth leads to an additional 0.7 years of education for spouses. The coefficient estimate is statistically significant at the 10% level. Among daughters the estimated coefficient has a larger magnitude, but the smaller sample makes for this outcome makes the coefficients for this comparison less precise. We find no statistically significant effect of parental wealth on spouse's occupational income scores.

6.4 Home ownership

Finally, we explore the effects of parental wealth on two proxies for wealth: home ownership and home value. Home ownership is defined as living in a home that is owned or mortgaged (i.e. not a renter) and being a household head or married to the household head. We find no statistically significant effects on home ownership, and the coefficient estimates are approximately zero. The estimated coefficients of log home value are quite large, but the small sample of homeowners makes it difficult to obtain a statistically significant effect.

7 Conclusion

We leverage geographic variation in the boundaries of oil fields in Oklahoma to estimate the long-run effects of parental wealth shocks on children' outcomes. Oklahoma Territory was settled quickly and homesteaders were unaware of the locations of oil fields when they settled. We compare the outcomes of children of homesteaders with oil on their land to those without oil. Local resource shocks can have both an individual effect (by making an individual family wealthier) and a broader community effect (by making the community as a whole wealthier). By comparing close neighbors, we net out the community-level effects of oil discoveries. Our estimates represent a lower bound of the effects of parental wealth because of attenuation from out-migration before oil discoveries.

Homesteaders responded to wealth shocks by retiring and moving to towns or cities in Oklahoma. Their children then attained additional schooling and were much more likely to complete high school. Sons were more likely to be in the labor force, and worked in higher paying occupations. Daughters were less likely to be married in young adulthood, and those who did marry appear to have had more educated spouses.

We consider two possible mechanisms of the effects of parental wealth on children's education, which are not mutually exclusive. One mechanism could be that parents' geographic mobility improved children's access to schooling in towns and cities. This is in line with previous evidence that wealth induces migration to better neighborhoods (Golosov et al., 2024) and of the effects of neighborhoods on children's long-run outcomes (Chetty et al., 2016). Another possible mechanism is that wealth reduced household labor demands, especially in farming households, enabling parents to exit the labor force and children to remain in school. The effects of wealth on education could have an indirect role in children's earnings, occupational choices, and migration and marriage patterns.

This study informs economic theory and public policy discussions focused on intergenerational mobility. We expand previous evidence on the causal effects of parental wealth by exploring distinct effects on daughters, including studying marriage outcomes. Additionally, our findings help reconcile studies on the effects of wealth on children's education. In contrast to Bleakley and Ferrie (2016), we find positive effects on education in a historical setting, likely because wealth shocks from oil are larger and more passive than wealth shocks from winning a land lottery. Our findings indicate that parental wealth can have meaningful impacts on children's economic opportunities in adulthood.

References

Ran Abramitzky, Leah Platt Boustan, and Tamar Matiashvili. Intergenerational Mobility over Two Centuries. *National Bureau of Economic Research Working Paper Series*, (33330), January 2025.

Philipp Ager, Leah Boustan, and Katherine Eriksson. The intergenerational effects of a large wealth shock: White southerners after the civil war. *American Economic Review*, 111:3767–3794, 11 2021.

Anna Aizer, Sungwoo Cho, Shari Eli, and Adriana Lleras-Muney. The impact of cash transfers to poor mothers on family structure and maternal well-being. *American Economic Journal: Applied Economics*, 16(2):492–529, April 2024. doi: 10.1257/app.20210816.

Aditya Aladangady. Housing wealth and consumption: Evidence from geographically-linked microdata. *American Economic Review*, 107(11):3415–46, November 2017. doi: 10.1257/aer.20150491.

Alex Albright, Jeremy Cook, James Feigenbaum, Laura Kincaide, Jason Long, and Nathan Nunn. After the burning: The economic effects of the 1921 tulsa race massacre. *SSRN Electronic Journal*, 2021. doi: 10.2139/ssrn.3880218.

Douglas W. Allen and Bryan Leonard. Rationing by racing and the oklahoma land rushes. *Journal of Institutional Economics*, 16, 2020. ISSN 17441382. doi: 10.1017/S1744137419000110.

Douglas W. Allen and Bryan Leonard. Property right acquisition and path dependence: Nineteenth-century land policy and modern economic outcomes. *Economic Journal*, 131, 2021. ISSN 14680297. doi: 10.1093/ej/ueab030.

Vellore Arthi. “the dust was long in settling”: Human capital and the lasting impact of the american dust bowl. *The Journal of Economic History*, 78:196–230, 2018.

William G. Axinn and Arland Thornton. The influence of parental resources on the timing of the transition to marriage. *Social Science Research*, 21, 1992. ISSN 0049089X. doi: 10.1016/0049-089X(92)90008-5.

Sarah Baird, Francisco H.G. Ferreira, Berk Özler, and Michael Woolcock. Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6, 2014. ISSN 19439407. doi: 10.1080/19439342.2014.890362.

Jacob Bastian and Katherine Michelmore. The long-term impact of the earned income tax credit on children’s education and employment outcomes. *Journal of Labor Economics*, 36, 2018. ISSN 0734306X. doi: 10.1086/697477.

Gary Becker, Scott Kominers, Kevin Murphy, and Jörg Spenkuch. A theory of intergenerational mobility. *Journal of Political Economy*, 126:S7–S25, 10 2018. doi: 10.1086/698759.

R.M. Bell and Daniel Mccaffrey. Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology*, 28:169–181, 01 2002.

Aymeric Bellon, J Anthony Cookson, Erik P Gilje, and Rawley Z Heimer. Personal wealth, self-employment, and business ownership. *The Review of Financial Studies*, 34(8):3935–3975, 04 2021. ISSN 0893-9454. doi: 10.1093/rfs/hhab044.

Dan Black, Terra McKinnish, and Seth Sanders. The economic impact of the coal boom and bust. *The Economic Journal*, 115:449–476, 4 2005a.

Dan A Black, Terra G McKinnish, and Seth G Sanders. Tight labor markets and the demand for education: Evidence from the coal boom and bust. *Industrial and Labor Relations Review*, 59, 10 2005b.

Sandra E. Black and Paul J. Devereux. *Recent developments in intergenerational mobility*, volume 4. 2011. doi: 10.1016/S0169-7218(11)02414-2.

Hoyt Bleakley and Joseph Ferrie. Shocking behavior: Random wealth in antebellum georgia and human capital across generations. *The Quarterly Journal of Economics*, pages 1455–1495, 2016.

Dan T. Boyd. Oklahoma oil and gas production: Its components and long-term outlook. *Oklahoma Geology Notes*, 65:4–23, 2005.

Dan T. Boyd. Oklahoma Geological Survey: Geological map 37 [dataset]. Oklahoma, OK: OGS, 2025. URL https://www.ou.edu/ogs/maps/geologicmaps.

John C Brown and David Cuberes. The Birth and Persistence of Cities: Evidence from the Oklahoma's First Fifty Years of Urban Growth, 2020.

Kasey Buckles, Adrian Haws, Joseph Price, and Haley E.B. Wilbert. Breakthroughs in Historical Record Linking Using Genealogical Data: The Census Tree Project. *NBER Working Paper Series*, 2024.

B. M. Bullard. *Digest of Oklahoma Oil and Gas Fields*, volume 40 of *Bulletin*. Oklahoma Geological Survey, 1928.

George Bulman, Robert Fairlie, Sarena Goodman, and Adam Isen. Parental resources and college attendance: Evidence from lottery wins. *American Economic Review*, 11:1201–1240, 4 2021.

Aline Bütikofer, Antonio Dalla-Zuanna, and Kjell G. Salvanes. Breaking the links: Natural resource booms and intergenerational mobility. *The Review of Economics and Statistics*, 107 (2):306–323, 03 2025. ISSN 0034-6535. doi: 10.1162/rest_a_01256.

A. Colin Cameron and Douglas L. Miller. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015. ISSN 0022-166X. doi: 10.3368/jhr.50.2.317.

David Cesarini, Erik Lindqvist, Robert Östling, and Björn Wallace. Wealth, health, and child development: Evidence from administrative data on swedish lottery players. *The Quarterly Journal of Economics*, 131(2):687–738, 02 2016. ISSN 0033-5533. doi: 10.1093/qje/qjw001.

David Cesarini, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Östling. The effect of wealth on individual and household labor supply: Evidence from swedish lotteries. *American Economic Review*, 107(12):3917–46, December 2017. doi: 10.1257/aer.20151589.

Kerwin Kofi Charles, Erik Hurst, and Alexandra Killewald. Marital sorting and parental wealth. *Demography*, 50:51–70, 2 2013. ISSN 15337790. doi: 10.1007/s13524-012-0144-6.

Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. How does your kindergarten classroom affect your earnings? evidence from project star. *Quarterly Journal of Economics*, 126, 2011. doi: 10.1093/qje/qjr041.

Raj Chetty, Nathaniel Hendren, and Lawrence F. Katz. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902, April 2016. doi: 10.1257/aer.20150572.

J. Anthony Cookson, Erik P. Gilje, and Rawley Z. Heimer. Shale shocked: Cash windfalls and household debt repayment. *Journal of Financial Economics*, 146(3):905–931, 2022. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2022.09.008.

Dora L. Costa. *The Evolution of Retirement: An American Economic History, 1880-1990*. National Bureau of Economic Research, 1 1998.

Cushing Independent. Two prospecting rigs up. *Cushing Independent*, page 1, January 1912.

James Cust and Steven Poelhekke. The local economic impacts of natural resource extraction, 10 2015. ISSN 19411359.

Jonathan M.V. Davis and Bhashkar Mazumder. The decline in intergenerational mobility after 1980. *The Review of Economics and Statistics*, pages 1–47, 02 2024. doi: 10.1162/rest_a_01413.

Oscar William Davison. Education at statehood. *Chronicles of Oklahoma*, 28:63–80, 1950.

N. Meltem Daysal, Michael F. Lovenheim, and David Wasser. The Intergenerational Transmission of Housing Wealth. Technical Report 22-21, University of Copenhagen. Department of Economics. The Center for Economic Behavior and Inequality (CEBI), September 2023.

Clément de Chaisemartin and Jaime Ramirez-Cuellar. At what level should one cluster standard errors in paired and small-strata experiments? *American Economic Journal: Applied Economics*, 16(1):193–212, January 2024. doi: 10.1257/app.20210252.

Department of Transportation. Technologies of oil and gas exploration. https://primis.phmsa.dot.gov/comm/technologies.htm, 2025. Accessed: 2025-06-30.

Christian Dippel, Dustin Frye, and Bryan Leonard. Bureaucratic discretion in policy implementation: evidence from the allotment era. *Public Choice*, 199, 2024. doi: 10.1007/s11127-022-01019-8.

Lasse Eika, Magne Mogstad, and Basit Zafar. Educational assortative mating and household income inequality. *Journal of Political Economy*, 127, 2019. doi: 10.1086/702018.

Mikael Elinder, Oscar Erixson, and Daniel Waldenström. Inheritance and wealth inequality: Evidence from population registers. *Journal of Public Economics*, 165, 2018. doi: 10.1016/j.jpubeco.2018.06.012.

James Feyrer, Erin T. Mansur, and Bruce Sacerdote. Geographic dispersion of economic shocks: Evidence from the fracking revolution. *American Economic Review*, 107:1313–1334, 4 2017. doi: 10.1257/aer.20151326.

Kenny A Franks. *The Oklahoma Petroleum Industry.* University of Oklahoma Press, Norman, 1980.

Tom K. Fulton and Haynie Stringer. The dawning of geophysical exploration. *Oil-Industry History*, 10(1), 2009. Volume 10, Number 1, 2009.

Claudia Goldin. *Understanding the gender gap: An economic history of American women.* National Bureau of Economic Research, 1990.

Claudia Goldin. America's graduation from high school: The evolution and spread of secondary schooling in the twentieth century. *The Journal of Economic History*, 58(2):345–374, 1998. doi: 10.1017/S0022050700020544.

Mikhail Golosov, Michael Graber, Magne Mogstad, and David Novgorodsky. How Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income*. *The Quarterly Journal of Economics*, 139(2):1321–1395, 5 2024. doi: 10.1093/qje/qjad053.

David Grann. *Killers of the Flower Moon: The Osage Murders and the Birth of the FBI.* Knopf Doubleday Publishing Group, 2018.

Katherine Hauck. Individual level and farm level homesteading characteristics: A descriptive overview. (4566300), 2023a. doi: 10.2139/ssrn.4566300.

Katherine Hauck. Crops vs. fences: Did homesteaders and purchasers follow different investment strategies in their land? *SSRN Electronic Journal*, 2023b. doi: 10.2139/ssrn.4566304.

Katherine Hauck and Tiemen Woutersen. Dynamic learning over beliefs about farming in the american west. September 2023.

Richard Hornbeck. The enduring impact of the american dust bowl: Short- and long-run adjustments to environmental catastrophe. *American Economic Review*, 102:1477–1507, 2012.

Richard Hornbeck. Dust bowl migrants: Environmental refugees and economic adaptation. *Journal of Economic History*, 83:645–675, 9 2023. doi: 10.1017/S0022050723000244.

Elisa Jácome, Ilyana Kuziemko, and Suresh Naidu. Mobility for all: Representative intergenerational mobility estimates over the twentieth century. *Journal of Political Economy*, 133(1): 306–354, 2025. doi: 10.1086/732527.

Matthijs Kalmijn. Intermarriage and homogamy: Causes, patterns, trends, 1998.

Seonghoon Kim and Kanghyock Koh. The effects of income on health: Evidence from lottery wins in singapore. *Journal of Health Economics*, 76:102414, 2021. doi: https://doi.org/10.1016/j.jhealeco.2020.102414.

Bryan Leonard and Brian Kogelmann. Does equality persist? evidence from the homestead act. *Journal of Political Institutions and Political Economy*, 3, 2022. ISSN 2689-4823. doi: 10.1561/113.00000059.

Bryan Leonard and Dominic P. Parker. Fragmented ownership and natural resource use: Evidence from the bakken. *Economic Journal*, 131, 2021. doi: 10.1093/ej/ueaa097.

Bryan Leonard, Dominic P. Parker, and Terry L. Anderson. Land quality, land rights, and indigenous poverty. *Journal of Development Economics*, 143, 2020. ISSN 03043878. doi: 10.1016/j.jdeveco.2019.102435.

Erik Lindqvist, Robert Östling, and David Cesarini. Long-Run Effects of Lottery Wealth on Psychological Well-Being. *The Review of Economic Studies*, 87(6):2703–2726, 2020.

Michael F. Lovenheim. The effect of liquid housing wealth on college enrollment. *Journal of Labor Economics*, 29, 2011. ISSN 0734306X. doi: 10.1086/660775.

Michael F. Lovenheim and Kevin J. Mumford. Do family wealth shocks affect fertility choices evidence from the housing market. *Review of Economics and Statistics*, 95, 2013. ISSN 00346535. doi: 10.1162/REST_a_00266.

Michael F. Lovenheim and C. Lockwood Reynolds. The effect of housing wealth on college choice: Evidence from the housing boom. *Journal of Human Resources*, 48, 2013. doi: 10.3368/jhr.48.1.1.

R. Mattheis and I. T. Raz. There's no such thing as free land: The homestead act and economic development. Unpublished manuscript, 2021.

B. R. Mitchell. *International Historical Statistics: The Americas, 1750–1988*. Palgrave Macmillan, Basingstoke and New York, 2 edition, 1993. ISBN 978-0333585153.

Magne Mogstad. The Human Capital Approach to Intergenerational Mobility. *Journal of Political Economy*, 125(6):1862–1868, 2017. doi: 10.1086/694633.

Amir Mousavi and Jeremy Edmund Clark. The effects of natural resources on human capital accumulation: A literature survey. *Journal of Economic Surveys*, 35:1073–1117, 9 2021. ISSN 14676419. doi: 10.1111/joes.12441.

Lucija Muehlenbachs, Elisheba Spiller, and Christopher Timmins. The housing market impacts of shale gas development. *American Economic Review*, 105:3633–3659, 12 2015. ISSN 00028282. doi: 10.1257/aer.20140079.

Arash Nekoei and David Seim. How do inheritances shape wealth inequality? theory and evidence from sweden. *Review of Economic Studies*, 90, 2023. doi: 10.1093/restud/rdac016.

Barack Obama. Remarks by the president on economic mobility. Technical report, The White House, Office of the Press Secretary, 12 2013. URL https://obamawhitehouse.archives.gov/the-press-office/2013/12/04/remarks-president-economic-mobility.

Edgar Wesley Owen. Rise of the mid-continent, 1911–1920. In *Trek of the Oil Finders: A History of Exploration for Petroleum*. American Association of Petroleum Geologists, 01 1975.

Frank Pega, Roman Pabayo, Claire Benny, Eun Young Lee, Stefan K. Lhachimi, and Sze Yan Liu. Unconditional cash transfers for reducing poverty and vulnerabilities: effect on use of health services and health outcomes in low- and middle-income countries, 2022. ISSN 14651858.

Joseph Price, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley. Combining family history and machine learning to link historical records: The census tree data set. *Explorations in Economic History*, 80, 4 2021.

Erin Ruel and Robert M. Hauser. Explaining the gender wealth gap. *Demography*, 50, 2013. ISSN 15337790. doi: 10.1007/s13524-012-0182-0.

Steven Ruggles, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler. IPUMS USA: Version 15.0 [dataset]. Minneapolis, MN: IPUMS, 2024a.

Steven Ruggles, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren. IPUMS ancestry full count data: Version 4.0 [dataset]. Minneapolis, MN: IPUMS, 2024b.

Sarada Sarada, Michael J. Andrews, and Nicolas L. Ziebarth. Changes in the demographics of american inventors, 1870–1940. *Explorations in Economic History*, 74:101275, 2019. ISSN 0014-4983. doi: https://doi.org/10.1016/j.eeh.2019.05.003.

John Steinbeck. *The Grapes of Wrath*. The Viking Press, 4 1939.

J Stuhler and F Biagi. A review of intergenerational mobility and its drivers. (KJ-NA-29366-EN-N (online)), 2018. doi: 10.2760/610312.

U.S. Department of the Interior. General Land Office (GLO) Records. https://glorecords.blm.gov, 2022. Accessed: 2022-12-02.

U.S. Energy Information Administration. U.S. Field Production of Crude Oil (Thousand Barrels). https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=mcrfpus1&f=a, 2025. Accessed 30 June 2025.

U.S. National Archives and Records Administration. Homestead act (1862). https://www.archives.gov/milestone-documents/homestead-act, 1862. Accessed: 2025-07-01.

U.S. National Archives and Records Administration. Dawes act (1887). https://www.archives.gov/milestone-documents/dawes-act, 1887. Accessed: 2025-07-01.

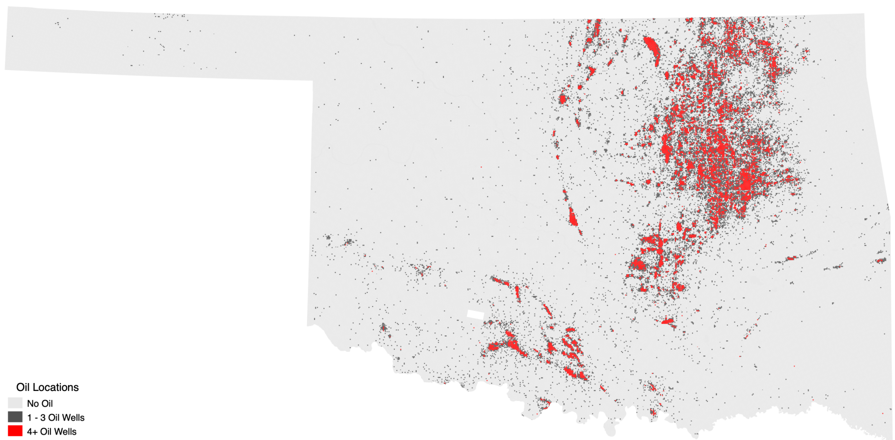
Micah Villarreal. Black gold: The effect of wealth on descendents of the enslaved. Draft forthcoming, 2025.

Zachary Ward. Intergenerational Mobility in American History: Accounting for Race and Measurement Error. *American Economic Review (forthcoming)*, 2023.

Julius Wilm, Robert K. Nelson, and Justin Madron. Homesteading. *American Panorama.*, 2025. URL https://dsl.richmond.edu/panorama/homesteading/. Interactive web map. Accessed 30 June 2025.

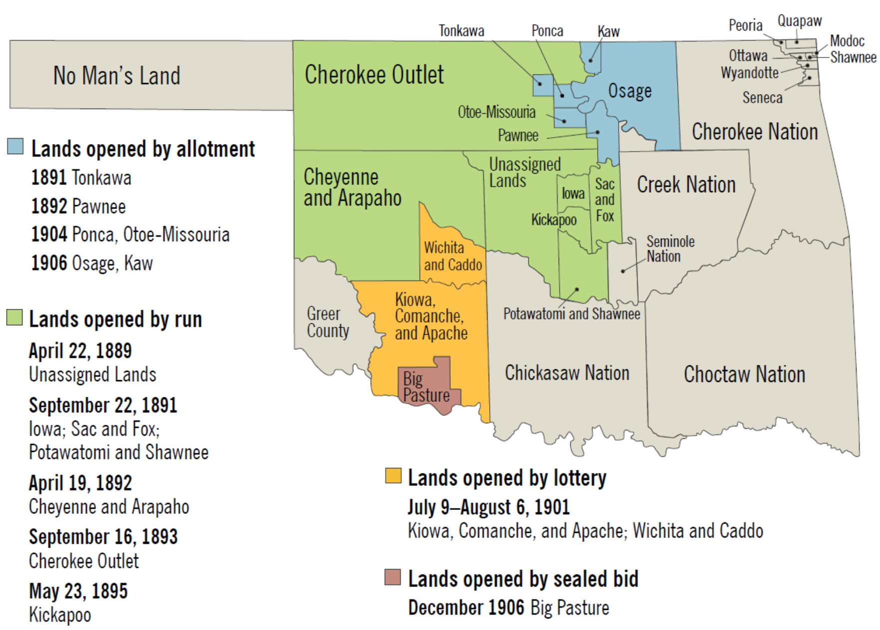
8 Tables and Figures

Figure 1: Oklahoma oil wells by 1940



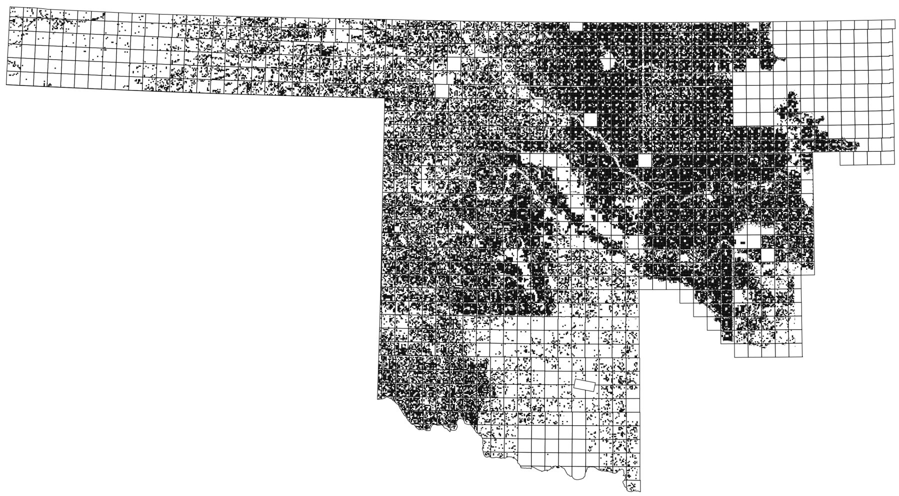
Notes: Locations of oil wells, from the Oklahoma Geological Survey Geological Map 37 by Dan T. Boyd. Wells are aggregated by aliquot at the quarter-section and lot level in the PLSS, with lots being irregular parcels of land shaped by geographic constraints. Fewer than 4 wells likely indicates wells that produce no oil.

Figure 2: Oklahoma Territory Land Openings



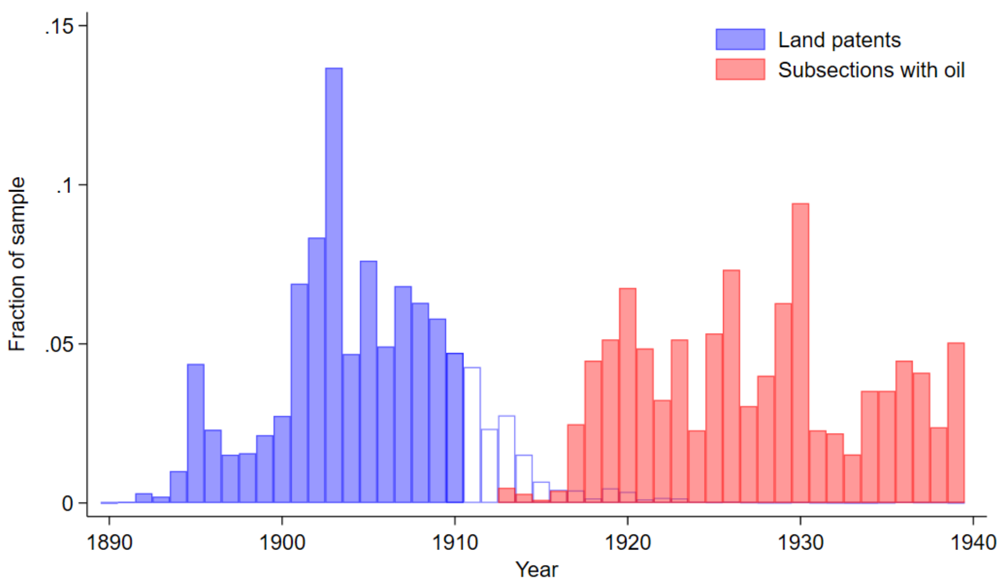
Notes: Map obtained from okhistory.org.

Figure 3: Federal Land Patents in Oklahoma Territory by 1910



Notes: Locations of land patents (titles) for homesteader entries in Oklahoma Territory by 1910. Data are obtained from the Bureau of Land Management.

Figure 4: Timing of land patents and oil discoveries in Oklahoma Territory



Notes: Distribution of land patents and oil discoveries in the former Oklahoma Territory. The unit of oil discoveries is aliquot parts in the Public Land Survey System, which are quarter-sections (quarter of a square mile) in most cases. Oil discoveries are counted as the first year when a subsection has at least 4 wells (a lower number likely indicates exploratory wells that produce no oil or gas.)

Figure 5: Sample construction process

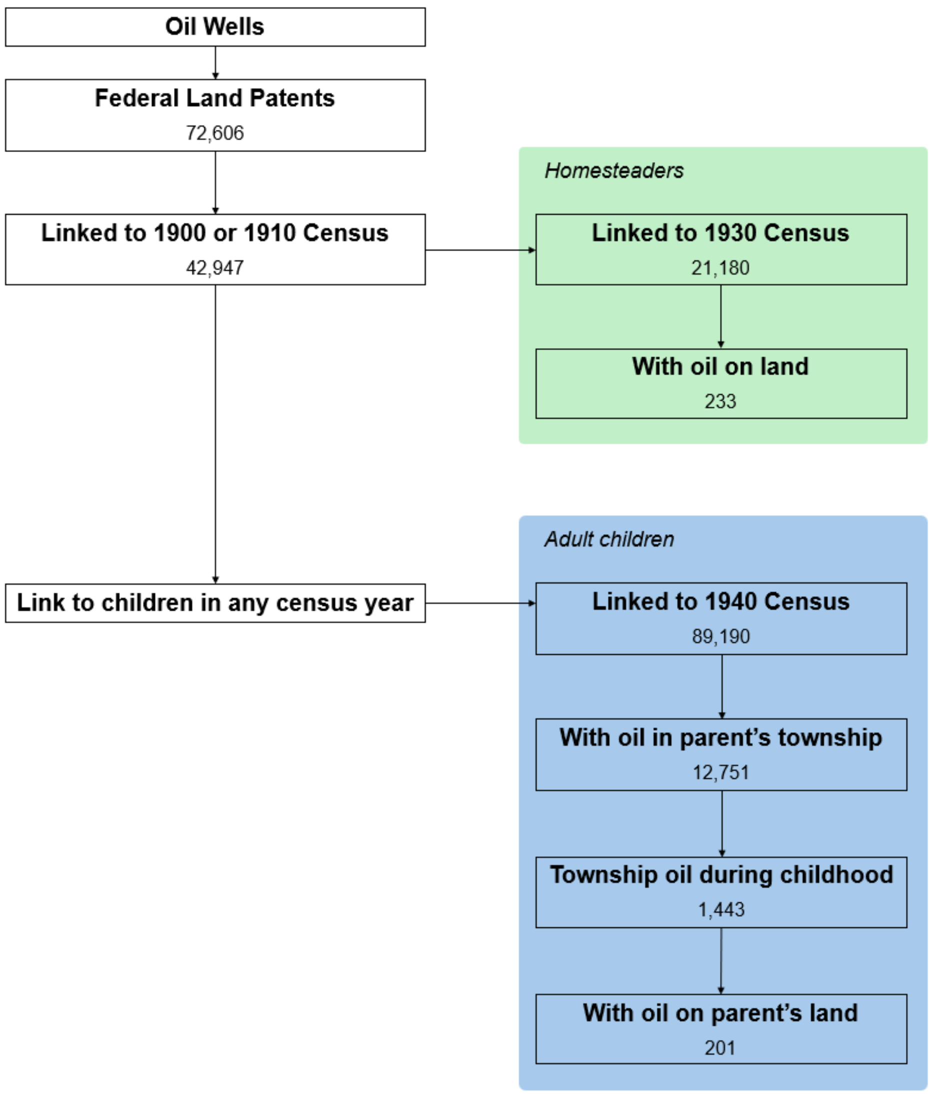
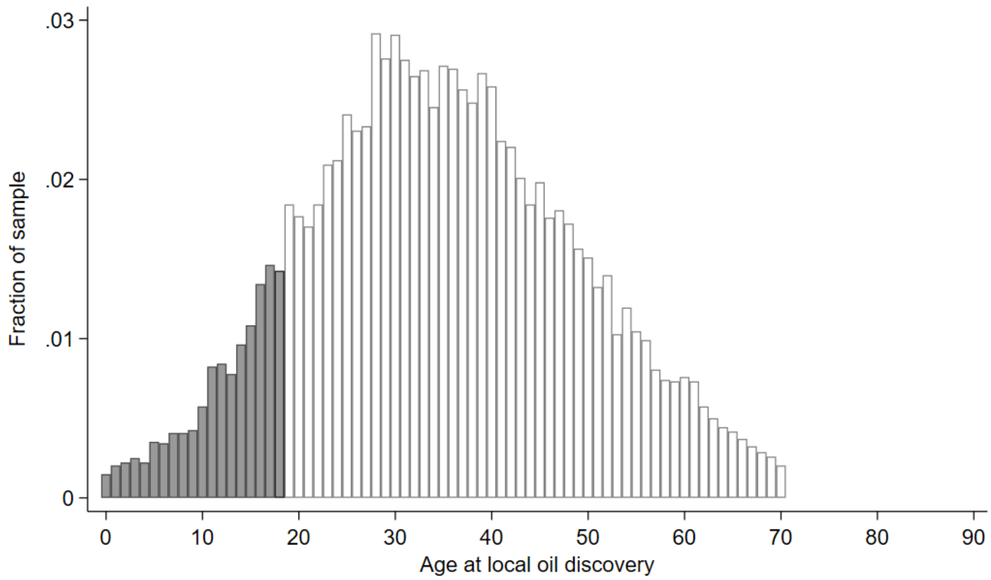


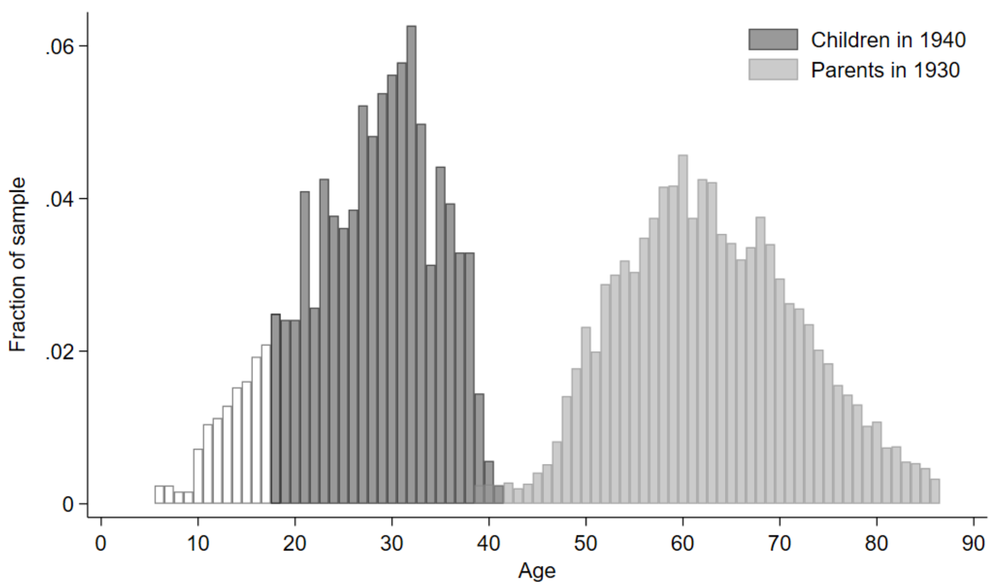
Figure 6: Age distributions

(a) Age at oil discovery, children



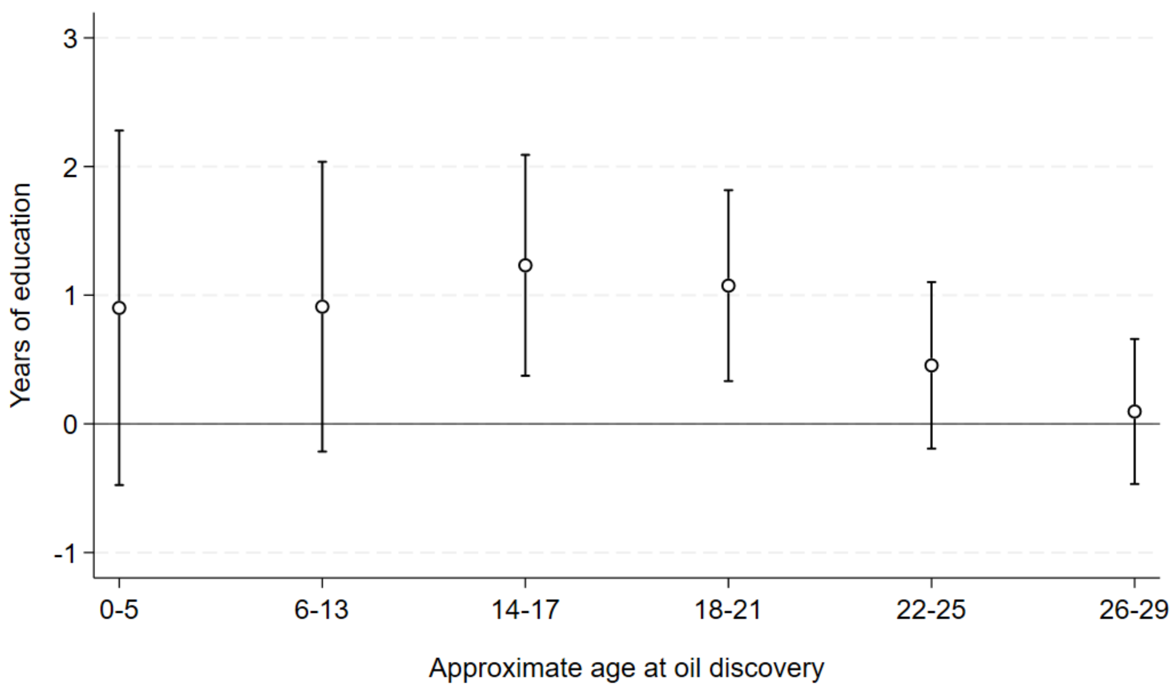
Notes: Distribution of ages at the median oil discovery date in their township. Sample includes children linked to the 1940 Census whose parent resided in a township that had with oil by 1940. The analysis is restricted to the ages with solid bars.

(b) Ages, children and parents



Notes: Distributions of ages in 1930 (parents) or 1940 (children). The children sample includes individuals linked to the 1940 Census who were under 18 when oil was discovered in their parent's township. The analysis is restricted to the ages with solid bars.

Figure 7: ITT effects of oil wealth on children's education, by age at oil discovery



Notes: Effects of having a parent with oil on their land, by age at time of the median oil discovery in the township. Completed years of education is measured in the 1940 U.S. Census among children at least 18 years old. Fixed effects are included for townships, age, and gender. We cluster standard errors at the township level.

Table 1: Mean characteristics of BLM land patents linked to a census record, by oil status

	Treated	Control	Within-township difference
Home owner	0.899	0.844	0.015
	[0.302]	[0.363]	(0.015)
Literate	0.954	0.950	0.002
	[0.209]	[0.217]	(0.008)
Occupational income score	13.928	13.368	-0.123
	[7.646]	[7.641]	(0.383)
Farmer	0.802	0.734	0.011
	[0.399]	[0.442]	(0.019)
White collar worker	0.050	0.041	0.000
	[0.217]	[0.198]	(0.009)
Craft worker	0.015	0.029	-0.011
	[0.123]	[0.167]	(0.007)
Manual worker	0.044	0.074	0.001
	[0.205]	[0.262]	(0.010)
Age	41.971	39.660	-0.700
	[12.758]	[13.720]	(0.686)
Female	0.084	0.094	0.012
	[0.277]	[0.292]	(0.014)
Black	0.025	0.029	-0.001
	[0.156]	[0.168]	(0.009)
Midwestern-born	0.610	0.556	-0.010
	[0.488]	[0.497]	(0.024)
Southern-born	0.219	0.284	-0.010
	[0.414]	[0.451]	(0.020)
Immigrant	0.080	0.087	0.012
	[0.272]	[0.282]	(0.014)
Married	0.762	0.728	-0.009
	[0.426]	[0.445]	(0.020)
Children	2.303	2.193	-0.006
	[2.305]	[2.242]	(0.127)
Observations	525	42,422	

Notes: Characteristics of individuals in the census who are linked to BLM land patents, with treatment based on the number of oil wells present by 1940. Mean values are reported, with standard deviations in brackets. Differences are computed as coefficients from a series of regressions with each characteristic as a dependent variable and an indicator for being treated as the key independent variable. We include township fixed effects and cluster standard errors at the township level (shown in parentheses). We use the value from the 1900 census, or from 1910 if the patent is only linked to 1910. The regression equations include a control variable for having unsure treatment status, as well as whether the patent is linked to the 1900 census. Ages are reported as of 1900.

Table 2: ITT effects of oil wealth on homesteaders

	In labor force	Farmer	Occupational income score	Urban	In Oklahoma
Treated	-0.116***	-0.066*	-2.231**	0.133***	-0.051
	(0.034)	(0.035)	(0.914)	(0.036)	(0.038)
Unsure	0.002	0.016	-0.453	-0.024	-0.027
	(0.018)	(0.022)	(0.490)	(0.019)	(0.020)
Observations	21,180	21,180	21,180	21,180	21,180
Mean, no oil	0.748	0.427	13.872	0.255	0.622

*$p < .1$, **$p < .05$, ***$p < .01$

Notes: Effects of having oil on homesteaders' land. "Treated" is defined as having a plot of land with at least 4 wells and "Unsure" is having land with at least one but fewer than 3 wells. Outcomes are measured in the 1930 U.S. Census. Fixed effects are included for townships, age, and gender. We cluster standard errors at the township level. Occupational income score is an occupation-based income measure, measured in hundreds of 1950 dollars.

Table 3: ITT effects of oil wealth on children's labor supply and income

	In labor force			Farmer			Occupational income score			Log annual earnings		
	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters	All
Treated	0.049**	-0.032	0.003	-0.023	0.005	-0.022	2.719**	-0.465	1.174	0.115	0.143	0.060
	(0.020)	(0.057)	(0.028)	(0.049)	(0.005)	(0.029)	(1.183)	(1.474)	(0.929)	(0.151)	(0.381)	(0.124)
Unsure	0.065***	-0.059	0.012	0.012	0.010	-0.007	1.279	-1.337	0.244	-0.017	-0.393	-0.026
	(0.020)	(0.055)	(0.022)	(0.043)	(0.010)	(0.027)	(1.260)	(1.254)	(0.873)	(0.084)	(0.457)	(0.082)
Observations	874	569	1,443	874	569	1,443	874	569	1,443	361	56	417
Mean, no oil	0.913	0.235	0.642	0.234	0.000	0.140	20.107	4.859	14.001	6.920	6.480	6.861

*$p < .1$, **$p < .05$, ***$p < .01$

Notes: Effects of having a parent with oil on their land. "Treated" is defined as having a plot of land with at least 4 wells and "Unsure" is having land with at least one but fewer than 3 wells. Restricted to children who were under age 18 when oil was discovered in the township and 18 or older in 1940. Fixed effects are included for townships and birth cohorts. We cluster standard errors at the township level. Columns with "All" include gender fixed effects. Outcomes are measured in the 1940 U.S. Census. Occupational income score is an occupation-based income measure, measured in hundreds of 1950 dollars. Effects on log annual earnings are restricted to the sample of wage workers.

Table 4: ITT effects of oil wealth on children's education and migration

	Education (years)			Completed high school			Urban			In Oklahoma		
	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters	All
Treated	1.088**	1.094***	1.069***	0.201***	0.110*	0.171***	0.062	0.071	0.074	-0.025	0.009	-0.011
	(0.498)	(0.401)	(0.393)	(0.073)	(0.062)	(0.053)	(0.060)	(0.075)	(0.050)	(0.075)	(0.074)	(0.062)
Unsure	0.250	0.579	0.414*	0.030	0.080	0.050	0.073	-0.011	0.054	-0.052	-0.123*	-0.088**
	(0.325)	(0.371)	(0.210)	(0.046)	(0.065)	(0.034)	(0.048)	(0.068)	(0.045)	(0.054)	(0.066)	(0.043)
Observations	850	554	1,404	874	569	1,443	874	569	1,443	874	569	1,443
Mean, no oil	10.291	10.823	10.505	0.411	0.525	0.457	0.336	0.387	0.356	0.552	0.591	0.567

*$p < .1$, **$p < .05$, ***$p < .01$

Notes: See notes for Table 3.

Table 5: ITT effects of oil wealth on children's marriage

	Ever married			Spouse's education (years)			Spouse's occupation score		
	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters	All
Treated	0.055	-0.110**	-0.012	0.383	0.850	0.656*	-1.450	1.254	0.198
	(0.047)	(0.052)	(0.034)	(0.434)	(0.647)	(0.381)	(1.354)	(2.139)	(1.169)
Unsure	-0.008	-0.070*	-0.025	-0.026	0.578	0.289	0.398	1.776	0.788
	(0.040)	(0.037)	(0.027)	(0.333)	(0.458)	(0.237)	(1.019)	(2.299)	(1.121)
Observations	874	569	1,443	560	388	948	570	397	967
Mean, no oil	0.703	0.768	0.729	10.661	9.959	10.372	2.583	23.712	11.314

*$p < .1$, **$p < .05$, ***$p < .01$

Notes: See notes for Table 3.

Table 6: ITT effects of oil wealth on home ownership

	Home owner			Log home value		
	Sons	Daughters	All	Sons	Daughters	All
Treated	-0.001	-0.002	0.004	0.588	0.415	0.349
	(0.055)	(0.074)	(0.042)	(0.356)	(0.479)	(0.283)
Unsure	0.020	-0.050	-0.009	-0.030	-0.497	-0.096
	(0.046)	(0.049)	(0.038)	(0.201)	(0.341)	(0.190)
Observations	874	569	1,443	224	160	384
Mean, no oil	0.260	0.293	0.273	6.772	6.949	6.847

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

Notes: See notes for Table 3. Home ownership is defined as living in a home that is owned or mortgaged (i.e. not a renter) and being a household head or married to the household head. Effects on log home value are restricted to the sample of home owners.

Online Appendix for "The Long-Run Effects of Parental Wealth Shocks on Children"

Cache Ellsworth Ian Fillmore Adrian Haws Joseph Price

A Data Sources

A.1 Federal Land Patents

We use the Bureau of Land Management (BLM) Federal Land Patents (U.S. Department of the Interior, 2022) to identify the name, precise location of residence, and patent (deed or title) date for over 170,000 land patents transferred from the federal government to individual homesteaders in Oklahoma Territory.[32] We focus on land patents that were filed between 1889 and 1910 in Oklahoma Territory. In addition, we also restrict our analysis to proved homesteads.

After filing for an initial homestead, there were three possible outcomes—prove, abandon, or commute the homestead. To successfully prove a homestead, individuals would've needed to apply for a patent after at least five years of "settlement and cultivation" (U.S. National Archives and Records Administration, 1862). According to Hauck (2023a), up to 45 percent of initial homestead claims were abandoned and returned to the federal government. Some individuals instead chose to commute their homesteads by purchasing them before the five years were completed.

We restrict our analysis to proved homesteads for two reasons. First, out-migration would exacerbate attenuation bias in our estimates of the causal effects of wealth. All individuals who abandoned their homesteads out-migrated, and Hauck (2023a) finds evidence that individuals who commuted their homesteads did so to resell the land. Second, homestead commuters received financial windfalls from reselling land in high-growth areas (Hauck, 2023a), introducing additional noise to our causal estimates. Additionally, our exclusion of commuted homesteads affects only a small portion of the BLM land patents.[33]

[32]Most of the land patents in Indian Territory (the eastern portion of present-day Oklahoma) were tribal allotments, which were quite different from the process of homesteading. Through the Dawes Act, the federal government broke up reservations and tribal lands into individual land allotments with the intention to forcibly assimilate Native Americans into white American culture through individual ownership and agricultural production (U.S. National Archives and Records Administration, 1887) In contrast to homesteads, allotments were often too small for farming. Additionally, many Native Americans were reluctant to take up agriculture or lacked the resources to do so. The Dawes Act also resulted in the dispossession of tribal lands because the government sold many allotments to non-Native purchasers.

[33]The land patents do not distinguish between individuals who directly purchased land from the government and those who commuted their homesteads. Approximately 29 percent of the non-allotment land patents we observe are reported as "cash entry", but it is likely that commutations form only a small fraction of cash entries. In Kansas, Hauck (2023a) finds that only 11 percent of cash entries were for commuted homesteads. We choose to omit purchasers because they were much wealthier than homesteaders and settled on land with different

The locations of land patents are identified using aliquots within a township and sections within the Public Land Survey System (PLSS). The PLSS defines a grid of six-mile by six-mile townships, each with 36 one-square mile sections (640 acres).[34] Typically, a homesteader received a quarter-section of land amounting to 160 acres. Thus, a section accommodates four homesteaders and their families, and a township accommodates 136 households.[35]

As shown in Figure B1, homesteaded land aliquots do not always align neatly within a single quarter section. It is rare for homesteads to extend across township boundaries—this occurs in only about 2% of cases. However, it is not uncommon for homesteads to span multiple sections which usually occurs when land is allocated to irregularly shaped lots. To address instances where a a homesteader is associated with multiple PLSS subsections and/or lots, we assign the subsection/lot that has the most oil wells. Specifically, we prioritize subsections/lots that have four or more oil wells. If four or more oil wells are not on any subsections/lots in the homesteading land we prioritize the subsection/lot that has any oil well. If there is no oil wells or multiple subsections with oil, we assign the subsection/lot with the largest acreage. This procedure ensures that each homesteader is linked to the land with oil or the largest subsection of their land.

A.2 Oil wells

For each homesteader, we identify whether oil was present on their land by linking the sections where they acquired land patents by 1910 to the locations of post-1910 oil wells. We use data from the Oklahoma Geological Survey (OGS) which provides precise locations and drilling dates for wells. We focus on oil wells in our analysis since gas wells didn't influence the wealth of homesteaders. Although gas wells existed before 1940, they were generally considered a nuisance or drilling hazard during the search for oil wells. Due to high pipeline construction costs and low natural gas prices, most gas was flared rather than captured. As a result, Oklahoma did

characteristics (Hauck, 2023a).

[34]Townships have a dual meaning in the PLSS. In addition to identifying a 36-square-mile area, the term "township" identifies the distance north or south of a referenced "baseline". Additionally, "range" identifies distances east and west of a referenced "meridian". For example, Township 3 North and Range 1 East locates the 36-square-mile township that is 12 miles north of a baseline and adjacent to the east of a meridian line.

[35]Two of the sections were set aside as "School Lands" and, at least initially, could only be leased, not homesteaded.

not significantly expand gas production until after 1940 (Boyd, 2005).

We define treatment status based on whether oil was likely present on a homesteader's land. Although the OGS provides locations and drilling dates for oil wells, it does not include information on production volumes. As a result, treating all drilled wells as indicators of oil discovery risks attenuation bias, since some wells may have been dry or yielded minimal production. To address this data limitation, we define a quarter section or lot as "treated" only if it contains four or more oil wells. Quarter sections with only one to three wells are marked and excluded from our final analysis. This threshold helps us more confidently identify land that had an economically meaningful oil discovery. To align oil discovery with our definition of treatment status, we assign the year of discovery as the year in which the fourth well was drilled within the subsection. The township discovery year is then the minimum year of discovery across all subsections within the township that meet the treatment threshold. This allows us to determine the earliest year at which economically meaningful oil activity began within the township.

Figure A1 color-codes each subsection of the Oklahoma City Oil Field based on OGS-reported oil activity, overlaid on a historical map that shows the precise locations of completed wells. In Figure A1a, we display treatment status by quarter section and lot without showing the OGS well locations. This allows for a clearer view of well placements as depicted in the historical map. In contrast, Figure A1b overlays the exact OGS well locations on the same historical map alongside our treatment classifications. This comparison provides a clear sense of how the OGS well data aligns with the historical map and how our treatment coding corresponds to actual drilling activity. Our threshold does a good job at identifying economically meaningful oil activity which effectively minimizes both false positives and false negatives. Overall, the strong correspondence between our treatment classification and the historical record supports the validity of our threshold.

As additional validation of our treatment classification and oil data, we provide a 1928 map of oil-producing fields in Figure A2. This historical map by Bullard (1928) shows the reported discoveries of oil fields rather than precise drill well locations. Despite the inherent challenges of using well data to infer oil presence at the subsection level, our well-based treatment measure offers the necessary level of granularity and aligns closely with this historical record.

Figure A1: Oklahoma City Oil Field Map by August 1930

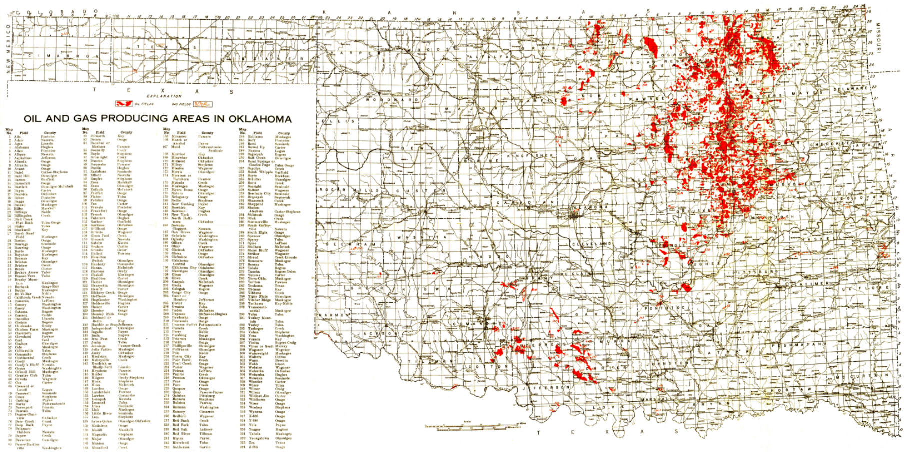
(a) Treatment | (b) Treatment with Oil Wells

A.3 Oil production volumes

Oil production values are not available in the OGS dataset so we supplement the data with another oil data source. We obtain data on oil production from the U.S. Geological Survey (USGS) which lists oil data wells by year in one-mile-by-one-mile cells and cumulative oil production through 1968 in two-mile-by-two-mile cells. Although this isn't a perfect measure of production, we hope to have an estimate of expected wealth for homesteaders.

We impute production for each homesteader as follows. First, for each two-mile grid cell, we divide cumulative 1968 oil production by the number of USGS wells in 1968, which gives us cumulative production per well in each two-mile cell. Then, for each of the four one-mile cells in a two-mile cell, we estimate cumulative production in 1940 by multiplying production per well

Figure A2: 1928 Oil Pools, from Digest of Oklahoma Oil and Gas Fields



Notes: Map obtained from Bullard (1928).

in 1968 (for the two-mile cell) by the number of wells in 1940 (in the one-mile cell). If a one-mile cell has fewer than five wells in 1940, we set its production to zero.

Because the one-mile USGS grid does not perfectly align with the one-mile PLSS sections, we assign each one-mile cell to the PLSS section that contains its center point. After this spatial alignment, we allocate one-fourth of the section's production to each homesteader in the section (reflecting the typical four homesteaders per section). This production is then assigned only to those homesteaders who are in the same PLSS section and meet the treatment threshold based on our OGS well method.

Finally, we convert imputed production into net present value of production royalties by applying a 12.5% royalty rate to the estimated oil volume and multiplying by the average oil price over the five years following the section's discovery year.

Table A1 presents summary statistics for imputed oil production and associated expected wealth or net present value of production royalties estimates among children of treated homesteaders. Column 1 shows the distribution of cumulative oil production per treated homesteader, with a mean of approximately 610,000 barrels and a very wide standard deviation. This variation reflects the wide distribution of oil production documented in Brown and Cuberes (2020). Columns 2 and 3 display the corresponding wealth estimates based on this production: first in 1940 USD, and then inflated to 2025 USD. On average, treated homesteaders gained an estimated \$90,270 in 1940 dollars, or approximately \$2,047,124 in today's terms, assuming a 12.5% royalty rate and average post-discovery oil prices over the next five years.

Although the distribution of oil wealth is highly left-skewed—with the top 10% of producers far outpacing the rest—the median treatment effect remains economically meaningful. The median estimated wealth is \$13,782 in 1940 dollars (or \$312,541 in 2025 USD), far exceeding the \$981 average annual income in 1940 USD (or \$21,356 2025 USD) for untreated children with positive income. These findings help verify with historical accounts (Franks, 1980) that oil discoveries were a transformative economic shock for many homesteading families. Still, it's important to note that these are cumulative production values as of 1968, and actual wealth realization may have been diminished by cash constraints or delays in extraction and payment.

Table A1: Oil and wealth estimates for total oil production

	Oil Production	Wealth Estimate (\$1940)	Wealth Estimate (\$2025)
Mean	609,838	90,270	2,047,124
SD	1,624,823	210,880	4,782,291
p10	3,042	618	14,008
p25	16,206	3,291	74,633
p50	66,164	13,782	312,541
p75	327,381	63,973	1,450,780
p90	1,382,797	232,732	5,277,850
N	742	742	742

Notes: Wealth estimates are sampled from children with treated sections and positive production in the USGS oil data (a small number of treated children have zero production reported in USGS). Production is imputed by allocating the cumulative 1968 oil production in every two-mile-by-two-mile cell equally among each of its one-mile-by-one-mile cells with 4 or more OGS wells by 1968. We assign each one-mile-by-one-mile cell to the closest section and divide the production amount by four (the typical number of homesteaders in a section). Column 1 includes all treated children with positive production and shows the estimated oil production in 1968. Columns 2 and 3 shows the expected wealth estimates in \$1940 and \$2025 USD respectively. For comparison, median annual earnings in 1940 are \$981 in 1940 USD and \$21,356 in 2025 USD among of our sample of untreated children with positive income.

B Record Linking Methodology

Our analysis of the intergenerational effects of wealth relies on constructing a panel dataset linking homesteaders to their children's long-run outcomes. We first develop a method to link land patent records to the U.S. Federal Census in 1900 and 1910. This step is critical for identifying family relationships, but it presents a difficult record linking challenge because of the sparse biographical information available in the land patents. We develop a strategy that leverages information on neighbors in both the land patents and the census. We then use individual census linkages from the Census Tree Project (Price et al., 2021; Buckles et al., 2024) to identify homesteaders' children and track children's long-run outcomes. The sample construction process is outlined in Figure 5.

Recent advances in historical record linking have successfully matched up to 80% of individuals across population-level U.S. censuses (Buckles et al., 2024). Linking methods mostly rely on a core set of personally-identifiable information (PII) available in historical census records, including names, birth years, and birth places. Yet two limitations result in persistent non-links and, to a small degree, inaccurate links. First, the set of PII is non-unique, especially because the census does not systematically collect granular birth dates, birthplaces, or middle names. Second, historical census records are prone to errors through the flawed processes of data reporting, enumeration, and digitization.

The information in Bureau of Land Management (BLM) Federal Land Patents (U.S. Department of the Interior, 2022) is much more sparse than census records. Land patents only include an individual's name and the location and date of the patent (equivalent to a "deed" or "title"), making it difficult to link patent holders to the census. Approximately 10% of a similarly sparse data source, inventor patents, have been matched to the census (Sarada et al. 2019).[36] Many recent studies have used historical land records such as land patents, but very few studies have attempted to link these to the census.[37]

We link 59% of land patents in our sample to the U.S. 1900 or 1910 Census by leverag-

[36]These contain only the inventor's name, residence place, and date of filing.

[37]Hauck (2023a) links initial land patents to the census, but does not provide the information needed to evaluate match rates. Villarreal (2025) links allotment records to the census with a similar match rate to ours, but is able to use information on family members contained in these records.

ing precise geographic information on neighboring homesteaders. By referencing the names of neighbors in the land patents and the census records, we are able to find more links and increase our confidence in their accuracy. We describe our linking process in the following sections.

B.1 Data pre-processing

We restrict the 1900 and 1910 Census records to only include adults, based on requirements of the The Homestead Act of 1862.[38] Additionally, we restrict the sample to Oklahoma residents because our linking strategy requires observing neighbors. We are unable to link individuals if they were not living on their homestead in the 1900 or 1910 Census, but this is favorable for our analysis because early out-migrants would further attenuate our causal estimates.

We process the land patents to extract useful information for linking. First, we map the patents into county boundaries at the time of the census, making it possible to restrict the set of candidate links to a closer geography. Second, we use naming patterns in the census to infer gender from names. If the gender is ambiguous, we consider possible links for both men and women. There were no gender restrictions in the Homestead Act, and approximately 14% of homesteaders in our sample had female names (see Table B1). We also clean information on current and former surnames, which are often reported for married or widowed women who obtain a land patent.

B.2 Generating candidate links

We select blocking rules to limit the set of candidate links and reject implausible links. There are over 72,000 land patents and approximately 218,000 adults in the 1900 Census, and the Cartesian product would result in 16 billion comparisons. We limit candidate links to those that match on county and gender (with an exception for ambiguous names), and have similar names.[39] After blocking, we retain approximately 56,000 candidate links to 1900, with most

[38]The Act applied to individuals at least 21 years old or who were the head of a household.

[39]We follow three steps to filter names. If a former surname is available, we use the former surname if linking to an earlier census and if the marital status in the census is single. First, we standardize names in the patent and census data (e.g. Wm is standardized to William). Second, we require that a candidate link match on the first letter or at least one bigram across records, for both first names and surnames. This step reduces the computational burden. Third, we require a Jaro-Winkler string similarity score of 0.87 or greater for both names.

patent records having at least one candidate link.[40] We also generate candidate links to the 1910 census, among patents with a date of 1905 or later.[41]

B.3 Using neighbors to de-duplicate links

The remaining candidate links contain many duplicates on the basis of names and county of residence. Land patents contain no additional information about individuals that would allow us to accurately de-duplicate these candidate links. Historical record linking algorithms often rely on variables such as birth year, birth place, or family members, none of which are available.

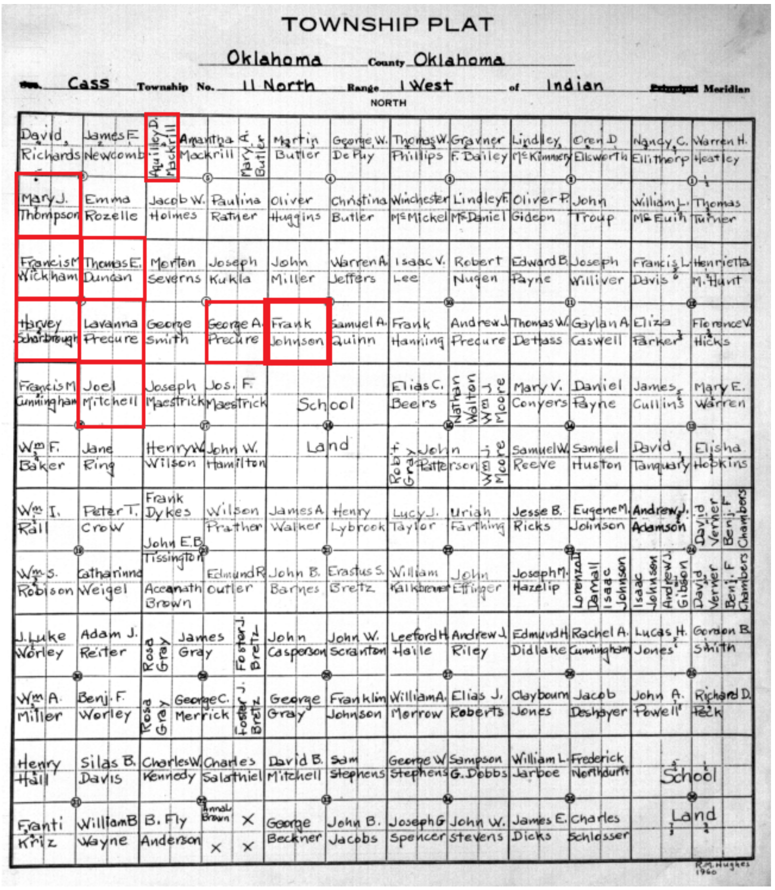
The geographic structure of the land patents provides the solution. See Figure B1 for a simplified illustration of our approach. Although there are multiple residents of Oklahoma County named "Frank Johnson", the land patents also provide the exact locations of Frank's neighbors. The highlighted neighbors in the plat map can be found on the same census enumeration sheet as Frank Johnson. Intuitively, the correct census record is the one that shares neighbors with the same name as the patent record. Although we do not observe house addresses in the census, neighbors are often enumerated on the same sheet because census workers typically walked door-to-door.

[40] Among the 31,000 patents between 1900 and 1904, 71% have a candidate link. There are on average 1.5 candidate links per patent retained.

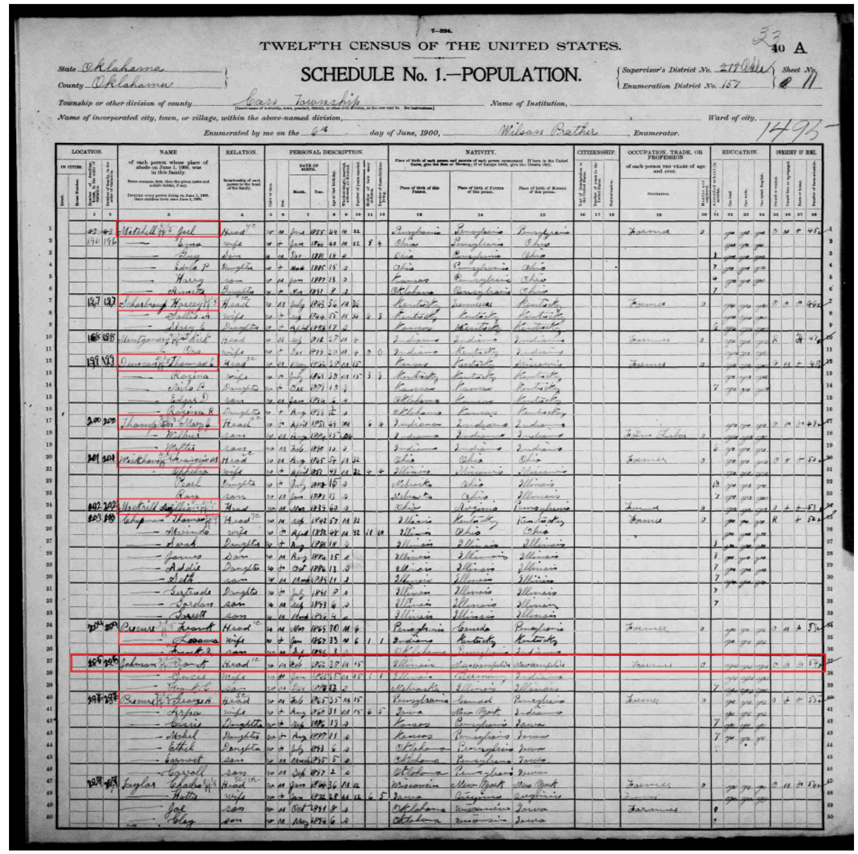
[41] Patents received in 1904 should correspond to settlement in 1899 or earlier, so these should all be in the 1900 Census. However, some patents in 1905 or later may have settled after 1900. If a patent is linked to census records in 1900 and 1910 that conflict with the link in the Census Tree, we retain the census link that is closest to the year of the patent.

Figure B1: Example of linking homesteaders to the U.S. Census

(a) Plat map



(b) Census page



Notes: Panel (a) shows the location of Frank Johnson and several neighbors in a township plat map in Oklahoma County, available from okhistory.org. Panel (b) shows where these neighbors can be co-located on a page of the 1900 U.S. Census in Oklahoma County, downloaded from FamilySearch.org.

We apply this idea by prioritizing candidate links with the highest number of neighbors linked between land patents and groupings of census households. We define neighboring land patents using 6-by-6 miles survey townships, which contain up to 144 homesteaders each. In the census, we define neighborhoods using groups of 50 consecutive households in the ordered census sheets. Among our candidate links, counties have 34 townships and 68 household groups on average. We de-duplicate candidate links by retaining those with the highest number of linked neighbors. A small number of duplicates remain; in some cases we can infer that these are duplicate patents, and we drop all remaining duplicates.[42] The remaining unique links to the 1900 Census have on average 17 linked neighbors.[43]

B.4 Results

We link 59% of homesteaders from a land patent to the 1900 or 1910 Census. This match rate is comparable to linkages across census records, despite the sparse information available in land records. Approximately one-third are not linked, which we attribute to data problems, migration, and surname changes among women. Census records often contain mis-spelled names or aliases that differ from the land patent (e.g. preferring to use the middle name). Some homesteaders also may have migrated before the time of the next census. Finally, surname changes due to marriage or remarriage after the land patent and before the census increase the challenge of linking women. Calculations from Table B1 shows that 41% of women are linked to a census record compared to 59% linked among the full sample. Table B1 also shows that homesteads that eventually have oil are no more likely to be linked than other homesteads within the same township.

Figure B2 plots the match rates by patent year. The match rate increases for patent years closer to a census, driven by out-migration. Former surnames in the patent data improve our ability to link women, but match rates are lower for women than for men.

[42] If one census record is matched to two patents in the same square-mile section, this likely indicates a duplicate land patent. If two census records are each matched to two patents in the same section, we find strong evidence that these are father/son pairs.

[43] Unique links to the 1910 Census have on average 8 linked neighbors. This number is lower because we attempt to link fewer patents to 1910.

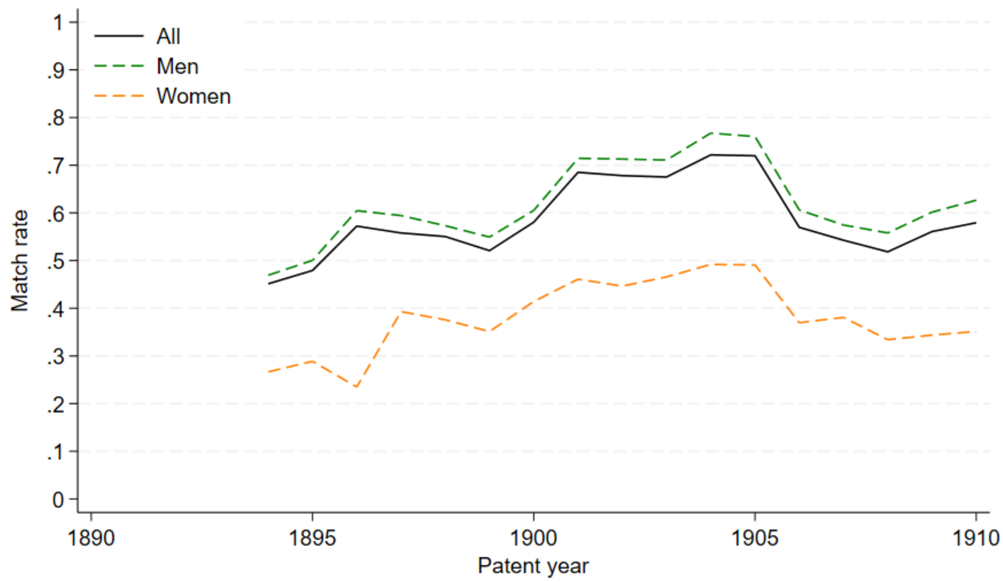
Table B1: Mean characteristics of BLM land patents, by linkage to a census record

	Linked	Unlinked	Within-township difference
Treated	0.012	0.010	0.000
	0.108]	[0.098]	(0.001)
Unsure	0.036	0.031	0.000
	[0.187]	[0.172]	(0.001)
Year of entry	1903.498	1903.387	0.275***
	[3.884]	[4.569]	(0.025)
Female	0.093	0.217	-0.127***
	[0.291]	[0.412]	(0.003)
Observations	44,559	28,047	

Notes: Mean characteristics of BLM land patents, with standard deviations in brackets. Differences are computed as coefficients from a series of regressions with each characteristic as a dependent variable and an indicator for being linked to a census as the key independent variable. We use township fixed effects and cluster the standard errors at the township level (shown in parentheses).

Figure B2: Fraction of Land Patents matched to the 1900 or 1910 census

(a) Fraction matched by gender



(b) Fraction matched by census year