A Mighty Toll:

Mine Accidents and the Long-Run Effect of Parental Loss

Ezra G. Goldstein*

December, 2023

Abstract

This paper estimates the causal effects of parental loss on a child's adulthood economic well-being by leveraging digitized records of nearly all early 20th-century U.S. mining accidents. I compare the outcomes of sons of fatal mining accident victims to those with fathers experiencing serious non-fatal accidents. Adult sons who lost their fathers when they were young experienced a 15 percent loss of income and had worse labor market outcomes. Examining families following the accident shows that widowed mothers were substantially more likely to take on the role of sole head of household and enter the labor market.

Keywords: Parental Absence; Parental Death; Orphans; Intergenerational Mobility
JEL Classification: J12, J62, N12

*eggoldstein@psu.edu Social Science Research Institute, Pennsylvania State University. I thank Ran Abramitzky, Jason Baron, Jon Denton-Schneider, Ellora Derenoncourt, Andreas Ferrara, Richard Hornbeck, Taylor Jaworski, Shawn Kantor, Carl Kitchens, Annalise Maillet, Peter Nencka, Sarah Quincy, Luke Rodgers, Tianyi Wang, Zach Ward, and Ariell Zimran. I also thank the many seminar and conference participants at Florida State University, University of Colorado Boulder, University of Notre Dame, the Southern Economic Association 2020 meetings, Cliometrics 2021, the 2021 NBER DAE Summer Institute, and the IPUMS Big Microdata Network Research Conference for their helpful comments and support. Special thanks to Ancestry.com and Gerald E. Sherard for providing the underlying data and making this research possible. All errors are the sole responsibility of the author.

“From the ranks of the true-hearted mates I have known
Death’s taken a mighty big toll
As, down in the mine, sweat-grimed they dig
For coal, coal, coal.”

— Pat Crorkin

“All for Coal.” *United Mine Workers Journal*, 1923

1 Introduction

It is difficult to overstate the formative role that parents play in a child’s development. Parental inputs such as care, attention, and household resources have been highlighted as some of the most crucial channels determining the future success of children (Cunha and Heckman, 2008; Dahl and Lochner, 2012; Del Boca et al., 2014). Beyond resources, parents act as a key source of human capital development. Through interactions with their parents, children develop the basic foundations for both cognitive and non-cognitive skills that play a direct role in their adulthood economic well-being (Seror, 2022). Moreover, the timing of parental investment matters, and investments during early ages of childhood generate the largest returns (Currie and Almond, 2011; Del Boca et al., 2014; Heckman et al., 2013).

Motivated by the critical role of parents, a large body of research has sought to uncover the consequences of disruptions to family structure and parental investments on a child’s development. An important strand of work highlights that growing up with an absent parent or separated household is associated with substantially worse measures of cognitive skills, mental health, and socioeconomic standing (Amato, 2000; Bloome, 2017; Corak, 2001; Gruber, 2004; Weinberg et al., 2019). Studies that examine this association tend to find modest causal effects, leading researchers to focus on perhaps the most severe disruption to a child’s family: the death of a parent (Corak, 2001; Lang and Zagorsky, 2001). The loss of a parent causes emotional distress and deprives orphaned children (those who have lost one or both parents) of an important channel for cognitive and non-cognitive development (Böckerman et al., 2023; Kalil et al., 2016). In addition, a parent’s death likely reduces the cumulative resources of the household and hinders investment in the bereaved children. As a result, the loss of a parent can have far-reaching implications for the bereaved child’s long-term outcomes.

Despite the importance of this topic, the long-term consequences of parental death during childhood and growing up without a parent are not well documented. Studying the effects of parental loss on a child’s adulthood economic well-being presents two distinct challenges. First, in order to examine a child’s adulthood outcomes, one must possess longitudinal data that links a child to their adulthood outcomes. The lack of long-term data has mostly limited studies of parental loss to focus on the health outcomes and educational attainment of children (Amato and Anthony, 2014; Case et al., 2004; Gertler et al., 2004; Gimenez et al., 2013; Weinberg et al., 2019). Second, even with sufficient data, one is left with the challenge of identifying a plausible counterfactual for those who experienced parental death during childhood. Given that the likelihood of experiencing parental loss is potentially related to factors that influence a child's adulthood success, it has been difficult to isolate the causal effects of a parent's death (McLanahan et al., 2013).

This paper estimates the effects of losing a father on a child's adulthood economic well-being. To do so, I combine a longitudinal dataset that links sons to their adulthood labor market outcomes and variation of parental loss due to historical mining accidents to estimate the causal effects of parental death. Specifically, I collect records from the early 20$^{\text{th}}$ century of more than 180,000 individual mining accident victims. By linking accident victims who were fathers to the full-count U.S. census, I am able to follow their sons to adulthood and directly measure wages, occupation, completed years of education, and other labor market outcomes.[1] Tracking the sons of accident victims enables the estimation of bereavement effects on intermediate outcomes, such as school attendance, as well as the outcomes of surviving mothers.

To identify the causal effects of the death of a father, I compare the adulthood outcomes of bereaved sons to sons whose fathers suffered a serious, but non-fatal accident. The validity of the identification strategy hinges on whether the loss of a father due to an accident was as good as random. I provide two key pieces of evidence that this was likely the case. First, I discuss the sudden nature of most mining accidents and highlight that, given their typical causes, the individual miner often had little control over the severity of a random accident. Second, I provide direct evidence that miners involved in non-fatal accidents were similar, on average, along several pre-accident characteristics relative to those who eventually perished.

Using a sample of sons of both fatal and non-fatal accident victims linked to their later-in-life outcomes in the 1940 census, I find that the early death of a father was associated with worsened economic well-being during adulthood. Those who lost their father during childhood experienced a roughly four percent lower annual wage income than those whose fathers survived. The estimate implies a loss of roughly 1.25 years of lifetime income for orphaned sons.

[1]Discussed below, linking individuals through the U.S. census requires matching on first and last names. Since women are likely to change their names upon marriage, it is generally not possible to link women throughout the U.S. census without additional information; the paper cannot estimate the effects of losing a father on daughters' outcomes. As a result, these data can only speak to the causal effects of parental death as it relates to fathers and sons.

Importantly, the overall effect of parental death is driven by a decline in wage income among those who were particularly young when they lost their father, supporting the broad phenomenon that shocks experienced during critical ages in childhood have far-reaching effects (Currie and Almond, 2011). Specifically, sons who lost their fathers when they were five years or younger experienced a 15 percent decline in wage income as adults. Testing additional measures of income, I find that sons who experienced the death of a father during early childhood were less likely to have an above-median income and fell 3.7 percentile points in the income distribution. I also find that orphaned sons had worse labor market outcomes more broadly. Sons who lost their fathers during early childhood were less likely to work in mining and instead worked in lower-skilled jobs, though the estimates are imprecise. In addition, sons who lost their fathers during early childhood were nearly 70 percent more likely to report being unemployed and 60 percent more likely to report being employed by a public relief program in 1940.[2]

While a detailed investigation of the precise mechanisms behind the loss of income experienced by bereaved sons is challenging due to the historical context and data limitations, I examine two broad channels suggested by the child development literature and the larger literature on orphans. First, I find that bereaved sons grew up in households with a broadly reduced capacity for parental investment. Widows were roughly four and a half times more likely to be listed as the head of household and about 60 percent more likely to be in the labor force. Given the strict, gendered division of household production among coal families of the era, the estimates suggest that the surviving parent had to assume a dual role of head of household and primary childcare provider. Second, I examine broad measures of human capital formation and show that sons who lost their fathers did not differ in their likelihood of attending school following an accident or their total educational attainment.

This paper makes two main contributions to the existing literature. First, this study contributes to the broad literature studying the long-run consequences of shocks experienced during early childhood. Notably, Currie and Almond (2011) highlights early childhood as a critical age for human capital development, and shocks experienced before age five can have persistent, long-term effects. A rich literature links disruptions to health (Grönqvist et al., 2020; Hoehn-Velasco, 2021), a parent's employment (Ruhm, 2004), and the home environment (Andrabi et al., 2021; Currie and Spatz Widom, 2010) to the future success of children across a varied set of contexts. A closely related vein of research similarly documents how shocks to family structure affect the well-being and development of children (Amato,

[2]During the depression era, public relief projects provided subsistence-level income for those without work. For many, work relief was seen as an alternative to unemployment and is largely viewed as a precursor to modern unemployment assistance (Fishback, 2017).

2000; McLanahan et al., 2013). Much of this literature tends to focus on spells of parental absence or divorced households, typically finding modest effects (Amato and Anthony, 2014; Gruber, 2004; Lang and Zagorsky, 2001). This paper adds to that literature by showing that the death of a parent—perhaps the most severe shock to a child's family environment—can have far-reaching effects, especially among those children who are particularly young when they experience the loss and among surviving spouses.

The findings of this paper also contribute to the literature studying orphans and their socioeconomic outcomes. The lack of available long-run data limits the empirical study of orphans' adulthood economic well-being, often constraining researchers to short-run survey data. As a result, most papers in this literature focus on orphaned children's early childhood and adolescent outcomes, such as physical well-being and educational attainment (Case and Ardington, 2006; Chen et al., 2009; Gertler et al., 2004; Gimenez et al., 2013; Weinberg et al., 2019). Only a few studies have been able to disentangle the effects of parental loss and unobserved, endogenous variation that also affects the success of children, often relying on exposure to disaster or disease to tease out causal estimates (Cas et al., 2014; Case and Ardington, 2006; Dribe et al., 2022; Gertler et al., 2004).[3] Much of the literature is mixed, typically finding at most modest effects of parental loss on a bereaved child's educational attainment (Beegle et al., 2010; Corak, 2001; Gertler et al., 2004; Kalil et al., 2016; Lang and Zagorsky, 2001). Likewise, the few studies that can follow orphans' adulthood outcomes tend to find similarly modest effects, and identifying plausible channels remains challenging (Adda et al., 2011; Dribe et al., 2022; Dupraz and Ferrara, 2021). This study adds to the existing literature by combining long-run, linked U.S. census data and a historical natural experiment to examine the adulthood economic well-being of orphaned children. The data allows for an analysis that directly estimates the effects of losing a father on adulthood wage income and identifies potential mechanisms. Specifically, the linked data allow me to examine the effects of death on educational attainment, occupational outcomes, and the outcomes of the surviving spouse.

The closest studies in this literature are Dupraz and Ferrara (2021), which examines the long-run socioeconomic outcomes of sons who lost their fathers in the U.S. Civil War, and Dribe et al. (2022), which examines parental loss and economic mobility in early twentieth-century Sweden. Their results show that bereaved children had lower occupational-based income scores and experienced occupational downgrading as adults, but given the sample period and data limitations, they cannot directly observe differences in average wage income,

[3]Another set of the literature attempts to identify the causal effects of parental death by taking advantage of variation in exposure to a parent while they were alive or via a sibling fixed-effects framework (Amato and Anthony, 2014; Kalil et al., 2016; Weinberg et al., 2019).

within occupation earnings, or completed years of schooling. This paper complements their work and the existing literature by studying the long-run outcomes of orphans in a setting where both wages and educational attainment are directly available. Moreover, that the average effects identified in this study are similar in magnitude to those in different settings is suggestive that the long-run effects of parental death may be modest.

2 Background

Coal Mining Families in the Early 20$^{\text{th}}$ Century

In the early 20$^{\text{th}}$ century, mining families and their communities represented a unique slice of American society. The industry offered relatively high wages (which compensated for the risks of the job) and attracted mobile workers, both native and foreign-born, to towns settled primarily to accommodate miners. Mining communities were also something of a melting pot; as the coal industry expanded from the late 19$^{\text{th}}$ to the early 20$^{\text{th}}$ centuries, so did employment opportunities for immigrant workers. By 1910, foreign-born workers, particularly those from southern and eastern Europe, constituted roughly half of the coal industry (Fishback, 1992).

To highlight the unique characteristics of mining families, I pool heads of households observed in the 1900-1920 decennial censuses in the four core states from which accident records were collected, Pennsylvania, Illinois, West Virginia, and Ohio. Panel A of Figure 1 then examines demographic characteristics, comparing coal industry families to other families in the early 20$^{\text{th}}$ century. Descendants from coal families were also distinct; in Panel B, I follow children through to adulthood in the 1940 census, comparing sons who grew up in coal industry families to others.[4]

Compared to other families of the era, coal families tended to be larger, living in households with 20 percent more children and were nearly 30 percent more likely to have a family size larger than the average family. Additionally, the early 20$^{\text{th}}$ century still saw a large share of American families on farms; 23 percent of heads of households worked in farming, and 22 percent lived on a farm. In contrast, just three percent of coal families lived on a farm, and more than half lived in an urban area. Mining families were also 30 percent less likely to own their homes, likely reflecting the high degree of families living in company housing. Families of coal miners were also nearly twice as likely to be foreign-born, with most immigrant families originating from Austria, Russia, Hungary, Italy, Germany, Poland, and the British Isles.

[4] I use a version of historical record linking techniques to match census records by state, full name, and birth year (Abramitzky et al., 2021). In Sections 3 and Appendix C.2, I discuss these methods in detail.

Following children through to adulthood in the 1940 census reveals that those from coal mining families also differed in meaningful ways. Adult men who grew up in a coal family were still less likely to live on a farm and more likely to live in an urban area. Sons of mining families also differed in their educational attainment and labor market outcomes. While 85 and 33 percent of children from other families earned at least an 8th grade and high school education, respectively, those from coal families were roughly 13 and 35 percent less likely to do so. In terms of their later-in-life economic well-being, sons from mining families were roughly ten percent less likely to earn an above-median wage.

Notably, the sons from coal mining families exhibited a remarkable propensity to enter the same industry as their fathers; sons from mining families entered the same industry as their fathers at nearly double the rate of other families. In fact, only agriculture surpassed coal mining in the rate of children entering their fathers' industry. The high rate of intergenerational transmission of occupation underscores the specific setting of coal families in the early 20th century, reflecting that, in some ways, coal mining was a family business (McGill, 1923). As a result, the potential disruption caused by losing a father may be particularly salient. Death not only came with emotional and financial consequences for a miner's children, but potentially the interruption of a channel for the intergenerational transmission of skills, tacit knowledge, and occupational connections.

The Dangers of Mining

Coal miners faced risks in nearly all aspects of their work. While workers in other industries faced similar threats of bodily harm, coal mining presented a greater threat to life. Fatalities per million man-hours in mining more than quadrupled levels seen in other industries (Fishback, 1992). Between 1900 and 1930, mine accidents claimed the lives of nearly 70,000 men and boys in the United States. Roughly three to four per thousand miners perished in an accident. In the early 1900s, mine disasters—large-scale accidents involving five or more fatalities—were relatively more common and killed about 400 miners annually. While disasters received the most attention, they accounted for just 16 percent of fatalities (Dix, 1978; Fishback, 1986).

Small-scale accidents—those that involved a single victim— comprised the majority of incidents, and about half of all mining accidents were caused by roof collapses and falling rocks. The typical accident occurred when a roof collapsed in part or entirely. Miners had to develop a sense for when a collapse would occur. An experienced miner could gauge the likelihood of a collapse by the sounds of the roof or the splintering of the timbers used to support it (Brophy, 1964). Still, miners were paid as a function of their output, so the prospect of extracting as much coal as possible tempted even the most experienced miners to overstay their welcome. When a roof did fall, the miner had little time to react before being struck by falling rocks or entombed by a collapse.

Miners also faced significant risks while setting charges to dislodge coal. In the early 1900s, miners constructed their own explosives using blasting powder and wax-like fuses, which were prone to fault. Accidents of this nature occurred when explosives detonated unexpectedly or reacted with flammable gasses present within the mine. Other circumstances were far beyond the control of the individual miner. Mine cars filled with heavy coal occasionally crushed workers. Faulty brakes, clothes being caught between wheels, or speeding drivers were often at fault (Fishback, 1992). Transportation equipment failure also resulted in fatalities. For instance, when the pulley system that raised and lowered the mine cage—the rudimentary elevator that transported workers throughout the mine—failed, miners plummeted down the mine shaft.

While fatal accidents represented the most severe outcome, non-fatal accidents were far more common. For instance, in 1930 there were roughly 1.5 times as many permanently disabling accidents as fatal accidents. Less serious incidents were even more commonplace; there were nearly 42 times as many accidents that caused miners to lose one or more days of work as there were fatal accidents (Adams et al., 1932). For these miners who survived, their injuries ranged from inconvenient to debilitating. In a survey of roughly 2,500 non-fatal mining accidents in the early 20^{th} century, about half of the injuries were less serious cuts, bruises, or burns (Roberts, 1904). Another 40 percent resulted in serious injuries, most typically fractures. The remaining non-fatal accidents resulted in debilitating and often disabling injuries, such as loss of limb, sight, or spinal damage.

Mine operators and supervisors provided little safety guidance to individual miners, who often worked alone or in pairs, sequestered within their workplaces (Fishback, 1992). Miners employed their own methods of accident prevention; For instance, they decided how often to bolster the support of the coal roof with timber in order to prevent collapses and how large of a blast to set in dislodging the coal from its seams. However, even when taking proper precautions, unavoidable and natural risks remained. A careless miner could quickly fall victim to a workplace accident that often resulted in death. Even still, careful miners were at risk of injury—or worse—due to negligence on the part of another worker, mine operators, or the sudden nature of accidents.

Accident Compensation

The most direct form of financial restitution to accident victims was compensation from employers. Prior to workers' compensation laws, eligibility for compensation was based on claiming employer negligence—that the accident was caused by an employer's failure to exercise "due care," for instance, failure to enforce safety in the mine or assignment of unqualified workers to dangerous tasks (Fishback, 1992; Fishback and Kantor, 2000). Consequently, many accident victims and their families went uncompensated.

In the eastern coal region of Pennsylvania, for instance, coal operators of the time claimed to wholly support the families of fatal accident victims. Yet, state inspectors reported that over 50 percent of accidents were claimed by operators to be due to carelessness and thus entitled families to limited or no benefits (Roberts, 1904). In a contemporaneous study of 137 fatal accidents in Pennsylvania, roughly 40 percent of families had been compensated and those families received compensation equivalent to the deceased's annual wage (Conyngton, 1917). Another case study examining records from a coal firm in Virginia indicated that the firm compensated three-quarters of families of fatal accident victims and half of families of serious, non-fatal accident victims (Fishback, 1992). For those incidents where employers deemed the worker at fault, miners (and their families) typically received no compensation. The remaining opportunities for compensation rested upon the generosity of local benefit funds, fraternal organizations, or success with legal action.

The enactment of workers' compensation laws in the mid-1910s marked a significant improvement, offering more substantial relief in terms of both the percentage of victims compensated and the amount of compensation they received (Fishback and Kantor, 2000). Under these laws, employers were obligated to provide compensation to injured workers or their families in the event of an accident. The compensation amount, predefined by officials, applied to any accident occurring during employment (Fishback, 1992). In Pennsylvania, for instance, employer payments to surviving families of fatal accident victims increased to about two to three times the average annual income of the deceased (Fishback, 1992; Fishback and Kantor, 1998). Compensation to surviving victims increased similarly; those who survived a mine accident received medical treatment for injuries and were compensated for lost wages at a rate equivalent to two-thirds of their regular pay.

Beyond employer compensation, miners and their families insured against the risk of injury and death by participating in local establishment funds supported mostly by workers (Fishback, 1992; U.S. Commissioner of Labor, 1908). Injured workers who participated in benefit funds were paid a portion of their earnings for up to six months or until they were able to return to work (Fishback, 1992). Temporary benefits to the injured and their families rarely exceeded more than one-third of typical wages for roughly up to six months. Relief funds typically provided greater relief to bereaved families; the families of fatal accident victims were entitled to roughly 57 percent of the deceased's annual earnings and additional compensation for each orphaned child. Still, death benefits for families rarely lasted beyond one year. Other local organizations also provided some financial relief to accident victims. For instance, in a Department of Labor report which highlighted the various benefits provided by local labor unions, local mine workers' unions often offered a modest, one-time payment to bereaved families, but few provided financial assistance to those temporarily injured (U.S. Commissioner of Labor, 1908).

During both eras of employer negligence and workers' compensation, the dependability and generosity of financial compensation clearly varied by the severity of an accident. Families of fatal accident victims were more consistent recipients of compensation, regardless of the funding source. Likewise, benefits disbursed to bereaved families generally approached or exceeded a worker's usual earnings, with bereaved families qualifying for benefits over longer periods of time compared to families of accident survivors. Regardless of compensation, the loss of life or limb resulted in significant financial turmoil for workers' families. A typical family in a mining town followed a strict division of labor where the husband took on the role of breadwinner, and household production typically belonged to the wife and children (Giesen, 1995; Roberts, 1904). In the event of an accident, the limited labor market opportunities for women meant that there was little opportunity for wives of injured miners to supplement income or replace it entirely. Even when such labor market opportunities were available, the wages received by women were low and the stigma of neglecting their children kept many from seeking employment (Lantz and McCrary, 1958; Wilson, 1922). Moreover, the rate of intergenerational transmission of occupations was high in the mining industry; when sons became working age they frequently joined their fathers in the mine (McGill, 1923). Thus, the loss of a father due to a mining accident presented not only financial and emotional hardship for families, but reduced economic opportunities on behalf of the children.

Altogether, the mechanisms and amounts of compensation available to accident victims in their families are starkly different than what is available in the modern social welfare system. Where both families of fatal and non-fatal accident victims faced an environment of uncertain financial relief in the early 20th century, today's nationwide life insurance companies and expanded public programs such as Social Security Disability Insurance and Social Security Survivors Benefits provide far more reliable and generous payments to families.[5] In light of

[5]For instance, establishment benefit funds often did not provide additional compensation per child of the deceased, and, when available, provided only modest additional support (U.S. Commissioner of Labor, 1908). In contrast, modern Social Security Survivors Benefits can be claimed by spouses caring for bereaved children who are under 16 and surviving children under 18. See, for example, Social Security Administration, *Planning for Your Survivors*. Accessed at: https://www.ssa.gov/

these differences, the estimated effects of losing a parent can be viewed as coming from a context with a particularly weak set of welfare programs, and are more likely to represent the net effects of parental death in the absence of a strong social safety net.

3 Data

In order to examine the effects of parental death on a child's adulthood economic well-being, I combine linked U.S. census data with individual mining accident records. In this section, I describe each source of data, present summary statistics, and discuss the construction of the main analysis sample.

3.1 Census Data

To study long-run outcomes, I rely on individual-level records from the U.S. census between 1900 and 1940 (Ruggles et al., 2021).[6] I utilize the 1900, 1910, and 1920 decennial census years to identify accident victims who were fathers at the time of enumeration. From these fathers, I identify the sons of accident victims and capture information about each household, such as the characteristics of the father and their family, and the county of residence. Specifically, I collect information about the parents' nativity, literacy, occupation, occupation-based income score, the number of children in the household, and whether or not the household contained extended family members.

To study the adulthood labor market outcomes of sons, I link forward individuals to the 1940 census. Importantly, the 1940 census marks the first decennial census year where all working-age respondents were directly asked about labor market outcomes, including wage income, and educational attainment. This information allows me to examine later-in-life economic well-being. In Section 5, I examine the intermediate outcomes of sons and mothers in the nearest census following an accident. To do so, I rely on the sons' responses in the 1910-1930 censuses. For instance, for the son of an accident victim identified in the 1900 census, I match that son forward ten years to the 1910 census. I then examine whether or not the sons reported attending school and their labor force participation, and also changes to the household and the labor market outcomes of surviving widows.

3.2 Accident Records

In response to growing concerns over dangerous conditions in the mining industry and several horrific disasters during the late 19$^{\text{th}}$ century, many states organized agencies to maintain

[6] In Section C.3, I discuss the data and the construction of census variables.

mine safety and report accidents. Each agency published an annual report to their state's legislature that compiled production statistics, provided updates on mine safety, and recorded the summaries of regional mine inspectors and officials. In some states, inspectors submitted the records of all accidents at each mine in their inspection region.

I collect these individual accident records from Pennsylvania, Illinois, West Virginia, and Ohio between 1900 and 1929.[7] I focus the analysis on these four states for two reasons. First, the records from these states are the most complete and consistently reported. Contemporaneous accident records from other states often report only fatal accidents, omit the age of the victims (making it impossible to link records to the census), or contain large gaps between issues. Second, these states were among the top producers of coal during the early 20th century. As a result, these states comprise the majority of mining accidents of the era.[8]

The accident records are bifurcated by the severity of injury to the victim; victims either suffered a fatal injury or one that was serious, but non-fatal. As discussed above, victims who perished as a result of an accident were often killed instantly and rarely survived past the accident date. Non-fatal accidents were typically reported if the accident resulted in 30 days or more of missed work. Figure A.1 highlights an example of the individual accident records published by state agencies. The excerpt lists fatal accident victims in Pennsylvania's Bituminous Coal Region during 1909. Importantly, the digitized accident records contain a number of useful features, including the names of victims, the dates of their accidents, and their ages.[9] From these reports, I collect any accident that lists the victim's full name and age at the time of their accident. I further discuss the construction of the individual accident dataset in Appendix Section C.1.

Altogether, the data contain 183,536 individual victims, 47,289 of whom perished. Table A.1 briefly summarizes the accident records. The average age at the time of the accident was roughly 34, though some boys and elderly are listed in the accident records. The majority of records come from Pennsylvania (64 percent), followed by Illinois (23 percent), West Virginia (11 percent), and Ohio (three percent). Panel A of Figure 2 summarizes the typical accident

[7] The records were assembled, digitized, and made available online by Gerald E. Sherard. I retrieved digitized accident records from the Russell L. & Lyn Wood Mining History Archive at the Colorado School of Mines and the Pennsylvania State Archives, accessed at https://libguides.mines.edu/ and http://www.phmc.state.pa.us/, respectively.

[8] I estimate that the assembled records account for approximately three-quarters of all fatal coal mining accidents during the early 20th century. Records of the U.S. Mine Safety and Health Administration show that between 1900 and 1915, a total of 33,053 miners perished in accidents within the United States. During the same period and throughout the four states covered in this study, I document 24,999 fatalities.

[9] Additionally, the records typically include a brief description of the cause of the accident. The descriptions of causes are most often short, limited phrases (e.g., fall of rock, from explosion, caught between mine cars) that occasionally provide some detail of the incident. Descriptions of the nature and causes of accidents were only digitized for Pennsylvania's records.

cause. About 80 percent of all accidents were caused by a roof collapse, the fall of rocks, or were due to faulty mine cars or other machinery. The likelihood of fatality varied by the type of accident; falling rocks and premature explosions were more often fatal than accidents involving mine cars and machinery or falls. While the cause of accidents varied by fatality, Panels B and C of Figure 2 show the distributions of birth year and age at the time of accident did not.

3.3 Linked Sample

I construct an intergenerational dataset of the sons of mining accident victims in two steps. First, I use the 1900, 1910, and 1920 decennial censuses to link the records of accident victims to the nearest census prior to an accident. Second, I identify sons of mining accident victims and link them forward to the 1940 Census. To link individuals across historical records, I employ a version of the Ferrie (1996) algorithm developed by Abramitzky et al. (2021) that minimizes spelling discrepancies between sources. In Appendix Section C.2, I provide further details of the step-by-step linking procedure.

In the first step, linking accident records to the census yields 53,510 unique matches—a match rate of 29.1 percent. Of these linked records, I identify 19,772 accident victims who were listed as heads of household at the time of enumeration. From these accident victims, I identify 16,279 native-born sons living in the accident victims' households to link forward to their adulthood outcomes in the 1940 census.[10] The second step results in a total of 9,346 adult sons of mining accident victims aged 20 to 60 at the time of the 1940 census. Finally, I focus on the 5,624 sons of accident victims who were eighteen or younger at the time of their father's accident, since older children are less likely to be living with their father at the time of the accident.[11]

[10]Sons are identified via the IPUMS variable *RELATE*, which describes the relationship between respondents and the head of household, and enables me to find all children of the household head. I focus only on native-born sons for two reasons. First, I focus on sons since women customarily changed their surname after marriage and it is generally not possible to link daughters without possessing additional information about renaming post-marriage. Second, I focus only on native-born sons since nearly 98 percent of identified sons are born in their state of residence.

[11]Since sons are identified as children of mining accident victims in the nearest census prior to the accident, I cannot directly observe co-residence at the time of the accident. I focus on sons less than 18 years of age at the time of their father's accident since they are likely to have been living with their parents. For instance, in the 1900-1920 censuses, 98 percent of boys aged 0-5 lived with either parent. The share slowly declines through childhood: 96 percent for those aged 5-10, roughly 93 percent for those aged 10-15. By age 16, still 90 percent of boys lived with their parents, and the share begins to shrink, but most remain co-resident: 86 percent live with a parent by age 17, and 80 percent by age 18. As boys entered adulthood, the share declines rapidly to half by age 22. Since living with a parent during early adulthood may be correlated to unobserved differences in childhood quality, I focus on sons who were younger than 18 at the time of their father's accident.

Examining the linked sample of fathers, I compare those who were accident victims to other fathers living in Pennsylvania, Illinois, West Virginia, and Ohio. There are two key reasons to expect substantial differences between the two groups: (i) selection into coal mining and (ii) selection due to the linking procedure.[12] Table A.2 compares fathers involved in accidents to the universe of fathers residing in the four states during the 1900, 1910, and 1920 U.S. censuses. Columns 1-3 display the means (and standard deviation in brackets) for all fathers, fathers involved in accidents, and fathers involved in fatal accidents, respectively. Columns 4 and 5 show the differences in means (and standard errors in parentheses) between all fathers and those involved in accidents, and between those involved in fatal accidents, respectively. The table shows that fathers involved in accidents and those who perished were younger, had larger families, and were less likely to own their home or live on a farm. In addition, fathers identified as accident victims were substantially more likely to be foreign-born and employed as mine operatives.[13]

Simply comparing sons who lost their father in a mine accident to sons of individuals who were not involved in an accident would conflate these differences. In such an analysis, controlling for observable characteristics would do little to ease concerns that there may exist some unobserved factors that differ between those involved in mining accidents and not. Describing these differences highlights the need for a plausible counterfactual that can credibly identify cases where the loss of a father could be unrelated to the characteristics of an average victim. As I discuss in the next section, the identification strategy relies on comparisons *between* fatal and non-fatal accident victims, and not individuals who were not involved in an accident. I show below that fathers involved in serious, but non fatal accidents did not systematically differ from those who perished.

Finally, in Section 5, I examine intermediate channels that may help to explain the potential effects of losing a father during childhood. To explore these potential channels, I follow the sons of accident victims to the nearest census following their fathers' accidents. Specifically, I attempt to match the 16,279 sons of accident victims identified within the 1900-1920 decennial census years forward ten years. As an example, for the sons of accident victims identified in 1920, I attempt to track the children in the 1930 census.

[12] Linked samples are often selected samples. Historical records that are linked to unique observations tend to be over-represented by those who were relatively well off and under-represented by immigrants and Blacks (Bailey et al., 2020).

[13] While Table A.2 shows fathers linked with the accident records were substantially different from other fathers residing in the same state, it does not speak to potential selection into the linked sample. In Figure C.2, I compare the accident records linked to the census to those that failed to find a match. While the comparison is limited by the available characteristics in the accident records, I show minimal selection into linking. Differences between matched and not-matched records are small in magnitude and precisely estimated. Still, to account for potential biases due to selection into the linked sample, I compute inverse propensity weights in Section C.2 and show the paper's main conclusions are unchanged after weighting.

This approach identifies 7,812 sons of accident victims who were eighteen or younger at the time of their fathers' accidents and aged between 10 and 28 when observed in the subsequent census. Observing these sons in the following census enables an examination of changes in schooling attendance and labor force participation after their fathers' deaths. In addition, I examine broad changes to parental investment by focusing on the outcomes of mothers who lost their husbands in an accident. Directly linking wives of accident victims presents a specific challenge: women typically change their names upon marriage, and since women who lost their husbands potentially remarry, linking women directly is generally not possible without external sources (e.g., marriage certificates). Instead, I focus on the mothers who could still be found in the households of accident victims' sons following the accidents. Identifying wives of accident victims in this way yields a sample of nearly 5,800 mothers observed after their husbands' accidents.

4 Empirical Strategy and Results

To quantify the effect of losing a father on a son's adulthood economic well-being, I focus on the log of annual wage income reported in the 1940 Census and examine differences between sons of fatal and non-fatal accident victims. In this section, I describe the identification strategy, probe its assumptions, and present estimates of the causal effects of parental loss on income during adulthood, along with several estimates of heterogeneity and robustness checks.

4.1 Empirical Strategy

To estimate the impact of losing a father on sons' adulthood economic well-being, I estimate the following equation on the sample of sons of fatal and non-fatal accident victims:

$$Y_i = \alpha + \beta FatherFatal_i + \mathbf{X_i}'\omega + \varepsilon_i \tag{1}$$

where the outcome variable, Y_i, represents the adulthood outcome of a son i observed the 1940 census; $FatherFatal_i$ is a dummy variable equal to one for an individual whose father perished in a mining accident; $\mathbf{X_i}$ captures both individual characteristics in the 1940 census and family characteristics observed during childhood. The coefficient of interest is β, which represents the average difference in adulthood outcomes between sons whose fathers perished as a result of a mining accident and those whose fathers suffered a serious, but non-fatal accident.

The identifying assumption is that the outcomes of sons of non-fatal accident victims constitute a valid counterfactual for the outcomes of bereaved sons. In other words, the assumption is that fatality from an accident is as good as random and unrelated to factors that may also determine a son's future economic success. For example, one may worry that foreign-born miners were less likely to internalize safety concerns relayed by their English-speaking supervisors and were more likely to perish in an accident as a result. In this case, comparing sons of fatal mining accident victims to sons of non-fatal accident victims would conflate differences in their fathers' nativity, which may influence their child's labor market outcomes.

If fatality is systematically related to individual characteristics of the father, one would expect to see differences between those eventually involved in fatal and non-fatal accidents. However, given the nature of mining accidents, there is little reason to believe this would be the case. As mentioned above in Section 2, accidents typically involved premature explosions due to faulty fuses, mine car failures, and the sudden collapse of a mine roof. As a result, an individual miner had little control over the severity of the accident, often due to negligence on the part of another miner.

In an effort to probe the main identifying assumption, Figure 3 plots the standardized differences in means between fathers who perished in a mining accident and two comparison groups: all other fathers who resided in the same state and fathers involved in serious, but non-fatal mining accidents.[14] The first set of estimates (in gray) plots the standardized differences between fathers who perished and all other fathers in the same state. As previously highlighted—compared to all other fathers, those involved in fatal accidents were substantially different.

In contrast, the second set of estimates (in blue) plots standardized differences between fathers who died and those who survived. Differences between fathers who perished and those who survived an accident are small, precisely estimated, and largely statistically indistinguishable from zero. The few characteristics that are significantly different by fatality differ by at most six percent of a standard deviation. Table 1 directly presents the summary statistics and differences in means between fathers who perished in an accident and those who survived. The table shows that, on average, fathers who perished in a mining accident were nearly identical to those who survived.[15]

It is important to be clear about the counterfactual. As discussed above in Section 3, mine inspectors recorded non-fatal accident victims if the accident kept an individual out of

[14]I standardize the estimated differences using the mean and standard deviation for each comparison group, separately.

[15]I also compare sons of fatal and non-fatal accident victims during childhood in Table A.3 and show that sons who lost their fathers in an accident were similar before the accident as well.

work for at least 30 days. Thus, the comparison is between sons who lost their fathers and those whose fathers were out of work for a noteworthy period due to an injury. However, the injury's severity was not recorded by mine inspectors, and non-fatal injuries ranged from briefly to permanently disabling. Given that the counterfactual in this context contains potentially debilitating injuries, the estimates below could understate the effects of losing a father. How much depends on the degree to which non-fatal accidents affected the long-run outcomes of victims' sons. In Section B, I estimate the direct effects of injuries on accident victims and their sons' adulthood outcomes, and show that the long-run effects of injuries are likely small.

4.2 Parental Death on Economic Well-being

Table 2 reports the estimates of parental death on annual wage income in 1940. Each column presents estimates from five separate regressions that progressively include covariates leading to the preferred specification. The first column includes adulthood characteristics observed in the 1940 census: a quadratic polynomial for age as well as adulthood county of residence and initial census year fixed effects. Column 2 controls for the fathers' pre-accident characteristics, which include the father's age at the time of the accident, nativity, literacy, the size of their family, occupation-based income score, and the industry they worked.[16] Column 3 includes childhood characteristics consisting of the son's age at the time of the accident, birth order, the number of older siblings and sisters identified during childhood, and whether or not their childhood household included extended family (e.g., grandparents or aunts and uncles), as well as fixed effects for the county of residence during childhood. Column 4 includes birth-year cohort fixed effects, which restricts comparisons to sons born in the same year and within the same state. Finally, the preferred specification is presented in the final column and further includes accident-year cohort fixed effects, which restricts comparisons to those with the same time since their father's accident occurred. Standard errors are clustered at the 1940 county of residence (shown in parentheses), which allows for arbitrary correlation in the structure of the error term across individuals residing in the same county.

Taken together, the estimates show that, compared to sons of non-fatal accident victims, losing a father as a result of a fatal mining accident was associated with a modest decline in

[16] While fathers who were involved in a mining accident worked in the industry at the time of their accident, they may not have during the nearest prior census. For instance, roughly 51 percent of fathers reported working in the mining industry. The other most commonly reported industries were none or not classified (14 percent), commonly associated with general laborers, followed by agriculture (10 percent), the railroad industry (3 percent), and construction (2 percent). Similarly, fathers involved in mining accidents most commonly reported their occupation as mine laborers (46 percent), general laborers (8 percent), farmers (8 percent), and management or foremen (4 percent).

annual wage income. On average, orphaned sons experienced between a 2.5 and 4.5 percent decrease in annual wage income as adults, an effect that is statistically significant at the ten percent level. The effect corresponds to a loss of lifetime income (discounted to 1940) of roughly $1,380, or about 1.25 years of lost wages.[17]

How do the estimates compare to others in the literature? While it is difficult to compare estimates across settings, the average effect sizes thus far are consistent with those found in the few other studies of the causal effects of parental death on earnings and socioeconomic mobility. In a context closely related to this paper, Dupraz and Ferrara (2021) examine the effects of losing a father in the postbellum U.S. Specifically, they show that sons of soldiers killed in action during the Civil War experienced a two percent decline in a measure of occupation-based income, though the effects are larger when fathers' deaths are instrumented with the mortality rate of their regiment. In another historical setting, Dribe et al. (2022) use census data from Sweden during the early 20$^{\text{th}}$ century and show parental mortality, driven by the Spanish Flu, reduced children's likelihood of upward mobility by between six and nine percent. Interestingly, they show effects are strongest for children who lost a parent at an early age and for those who lost their mother. Even in the modern context, estimates of the long-run effects of parental death are comparable; Adda et al. (2011) examine the long-run effects of the loss of both mothers and fathers on children's adulthood outcomes in Sweden during the early-2000s by using comprehensive administrative data and find a six to seven percent loss in earnings during adulthood. Taken together, these few studies that examine the long-run effects of parental death reveal strikingly similar effect sizes observed across varied contexts and empirical designs. Moreover, the pattern of estimates is close in magnitude with the estimated average decline of between 2.5 and 4.5 percent in annual wage earnings identified in this study and suggests that the average causal effect of parental death on adulthood economic well-being is likely to be modest.

4.3 Heterogeneity

In this section, I examine heterogeneity of the estimated income loss by family and child characteristics and augment Eq. 1:

$$Y_i = \alpha + \gamma FatherFatal_i + \tau Z_i + \theta(FatherFatal_i \times Z_i) + \mathbf{X_i}'\omega + \varepsilon_i \tag{2}$$

[17] To calculate the value of lifetime income loss, I assume that sons of accident victims work for 50 years. Given that the average wage income for sons of non-fatal accident victims was roughly $1,100 in 1940, the estimated difference corresponds to $44. I assume the difference follows the annual rate of inflation throughout the sons' working lives. Using a discount rate of 6 percent, which corresponds to the real return of the S&P between 1940 and 1990, I calculate the 1940 present value of the decline in wage income is roughly $1,380.

where Y_i, $FatherFatal_i$, and $\mathbf{X_i}$ are defined as in Eq. 1; Z_i represents characteristics of the child or their household environment prior to their father's accident. For instance, when examining how the effects of losing a father may be affected by the son's age at the time of the accident, Z_i is an indicator equal to one if son i was younger than primary school age (five years or younger) at the time of their father's accident. In this specification, γ and $\gamma + \theta$ represent the effects of losing a father for two different groups depending on their values of Z_i. When Z_i is binary, following the previous example, γ estimates the effects of losing a father for sons older than primary school age, while $\gamma + \theta$ does the same for those who were particularly young at the time of their father's accident. Alternatively, when Z_i is continuous, $\gamma + \theta$ is evaluated at the mean of Z_i. As an example, examining how the effects vary by having an older sibling, γ estimates the effects of losing a father for sons without older siblings, and $\gamma + \theta$ does the same for those with two older siblings. I examine subgroup heterogeneity in Table 3.

Family Supports

The loss of a father in the early 20$^{\text{th}}$ century meant a substantial shock to the household. Coal families followed a gendered division of labor; fathers entered the mine to earn wages while mothers engaged in domestic production at home (Giesen, 1995; Roberts, 1904). As a result, losing a father likely disrupted both household resources and broad family investment, such as the time spent with children or material investments. Then, it is natural to examine whether or not the availability of family support, with either time or resources, may have mitigated the documented loss of wages during adulthood.

The first two columns examine how the availability of family support impacts the long-run effects of paternal loss. In Column 1, households are identified by whether or not the father lived with extended family, specifically whether their household included their parent or siblings (i.e., the child's grandparent or aunt and uncle). The first estimate in the column shows that sons who lived in households without any extended family experienced a wage loss in adulthood similar to the baseline estimates. The second estimate provides some evidence that available family support may mitigate the long-run effects of losing a father; while imprecise, the point estimate for those who lived with extended family when they lost their fathers is close to zero. Similarly, Column 2 compares those without older siblings and those in a household with older siblings. While the estimates are not precise, the exercise again provides some suggestive evidence that those in households with less available family support experienced larger long-run effects of losing a father.[18]

[18] I also examine differences in losing a father by order of birth. For instance, lower-ordered-born sons may have been more insulated from the loss of their fathers if higher-ordered-born sons replaced their father's

When support by other family members was not available, support from the community may have reduced the long-run negative effects of losing a father. Columns 3 and 4 examine how the effects of losing a father may differ based on the relative and absolute size of the mining community, respectively. To measure the relative and absolute size of the mining community, I collect the share of the childhood county working in the mining industry and the total number employed in mining, respectively. Neither the share nor total size of the mining community appears to modify the long-run effects of parental loss significantly; both in Columns 3 and 4, the effects of losing a father in a county with an average-sized mining community are similar to the effects in a county where the community was one standard deviation larger.

Losing a Father at an Early Age

The experience of losing a parent may differ substantially based on the child's age when they lost a parent (Berg et al., 2016; Dribe et al., 2022; Gertler et al., 2004; Guldin et al., 2015). On the one hand, young sons may suffer more from losing a parent if the loss corresponds to worsened care-taking in the home or reduced parental inputs and human capital transmission during a critical age of development (Heckman and Mosso, 2014; Kalil et al., 2016; Weinberg et al., 2019). On the other hand, children who were school-aged when they lost a parent may have had to forgo education or other labor market opportunities to support their families (Gertler et al., 2004).

The last column of Table 3 shows that sons who lost their fathers are at a relatively older age and experienced little differences in wage income as adults compared to sons who were similarly aged when their fathers were involved in a non-fatal accident; the estimate is small, roughly 1.5 percent, and not economically meaningful. Conversely, sons who were young when their fathers perished in a mining accident experienced a substantial loss of adulthood wage income. Compared to sons who were similarly aged when their fathers survived an accident, sons who lost their fathers when they were younger than primary school age experienced roughly a 16 percent decline in wage income as adults.

To further examine the effect of the loss of a father by age at the time of the accident, I consider the effect throughout childhood. Specifically, I flexibly examine how the effects vary by age at the time of the accident. To do so, I group sons into four age bins: when they were between (i) 0-4, (ii) 5-9, (iii) 10-14, and (iv) 15-18 at the time of their father's accident. I then estimate the following equation:

labor or parental time. I interact the loss of a father with birth order and calculate the total effects of losing a father when being the first born, second born, or third or higher order born. Table A.4 presents the estimates and shows no discernible pattern.

$$Y_i = \alpha + \sum_{b=1}^{4} AgeBin_{ib}\nu_b + FatherFatal_i \times (\sum_{b=1}^{4} AgeBin_{ib}\eta_b) + \mathbf{X_i}'\boldsymbol{\omega} + \varepsilon_i \tag{3}$$

Where $AgeBin_{ib}$ equals one if a son was in age group b at the time of their father's accident so that $\widehat{\eta_b}$ captures the estimated differences between bereaved sons and sons of non-fatal accident victims in age group b. Figure 4 plots the $\widehat{\eta_b}$s and presents the effects of losing a father on later-in-life wage income by the son's age at the time of their father's accident. The leftmost estimate suggests the youngest sons who lost their fathers experienced the most severe income loss as adults. The effect declines gradually and becomes close to zero among the group that was between ages 10 and 14 when they lost their father.[19]

Taken together, the results of this section suggest that the baseline estimates presented in Table 2 mask substantial heterogeneity. Specifically, the modest decline in adulthood earnings is driven by pronounced differences in adulthood wage income for those whose fathers perished during early childhood. Sons who lost their fathers at an early age experienced roughly 16 percent lower earnings as adults, compared to a minimal income loss among those who experienced the loss of their father when they were relatively older. Moreover, allowing the estimates to vary further by the son's age at the accident reveals a noticeable pattern: the estimated effects of parental loss are more severe for children who were deprived of their father through most of childhood, and the effect weakens with increased time spent with their father.

The pattern of these results suggests a pathway for the observed, long-run effects of losing a parent. Parental investment earlier during childhood has been highlighted as a key determinant of a child's future success (Cunha and Heckman, 2008; Del Boca et al., 2014). Furthermore, shocks experienced by children by age five that disrupt investment and human capital accumulation can have long-term consequences for their well-being Currie and Almond (2011). Changes to a child's health (Grönqvist et al., 2020; Hoehn-Velasco, 2021) and parent's economic outcomes (Ruhm, 2004) when children are particularly young, for example, have been shown to have pronounced effects on human capital development. That the long-run effects of losing a parent are more pronounced for younger children perhaps suggests that disruptions to the family follow similar patterns (Dribe et al., 2022; Kalil et al., 2016). For that reason, throughout the remainder of the paper, I explore heterogeneity in the effects of losing a father, presenting the average effects of bereavement as well as the effects among those who were five or younger when they lost their father. Moreover, in Section 5, I

[19] I rely on four age bins due to small sample concerns. Using four age groups, the number of bereaved sons in each group is 188, 408, 475, and 654, respectively. In Figure A.2, I show that the pattern of estimates remains unchanged even as bin sizes shrink, though sample sizes in each bin become smaller and the estimates become less precise.

examine proxies for parental investment and human capital formation, which may, in part, inform the long-run effects of parental loss.

Additional Measures of Income

The results thus far show that sons who lost their fathers, and in particular those who were young at the time of their father's death, experienced reduced economic wage income as adults. However, wage income is just one potential metric of economic well-being. In this section, I additionally estimate how losing a father affected sons' position in the income distribution and whether or not they received non-wage income.

To uncover how sons of fatal accident victims fared in the income distribution, I instead compute the median wage income of working-age men born in the four states covered in the accident records and construct an indicator variable for whether the adulthood wage income for sons of accident victims was above the median. I then re-estimate β from Eq. 1 and $\gamma + \theta$ from Eq. 2 to examine how losing a father, and losing a father at an early age affected the likelihood of having above-median wage income.

Column 1 of Table 4 displays the estimated differences. The estimates suggest that, compared to sons of non-fatal mining accident victims, orphaned sons were slightly less likely to have wage income above the median. However, the estimate is small (two percentage points and three percent relative to the control group mean) and insignificant. Panel B shows that the average masks differences among those relatively young at the time of the accident. Sons who lost their fathers at a relatively early age were six percentage points less likely to have an above-median income (approximately ten percent relative to the control group mean).

To further explore how parental death affected the sons' positions in the income distribution, I examine their rank in the income distribution. The second column of the table shows that bereaved sons were roughly one percentile point below those whose fathers survived an accident. Again, the average difference is small relative to the control group mean and statistically indistinguishable from zero. Panel B, however, suggests sons who lost their fathers during early childhood were three percentile points lower in the income distribution. Given that sons of non-fatal mining accident victims were, on average, in the 44^{th} percentile, this implies that orphaned sons fell to the 41^{st} percentile in the income distribution.

Thus far, all measures of economic well-being have relied on reported wage income during adulthood. Yet wages are only one component of total income. Notably, wage income (recorded in the 1940 census) does not include business or farm earnings, unemployment payments, or other forms of government and private compensation (Census Bureau, 1940).

To further examine how parental loss affected total income, I consider differences in non-wage income of orphaned sons.[20] In Column 4 of Table 4, I show that sons who lost their fathers in a mining accident were roughly 3.5 percentage points (25 percent relative to the control group mean) more likely to report having greater than $50 of non-wage income, on average. Panel B shows that sons who were particularly young at the time of their father's accident were similarly more likely to report non-wage income, although the estimated difference is imprecise.

The fact that sons who lost their fathers are more likely to report having non-wage income raises several questions. Given the scope of the category, it is difficult to pinpoint the precise source of increased non-wage income or its magnitude. One may wonder if the increased probability of non-wage income and reduced wage income reflects that sons who lost their fathers substituted away from wage-earning occupations and moved into occupations such as business or farm ownership. Alternatively, higher non-wage income may reflect an increased likelihood of unemployment compensation or government and private relief eligibility. Below, I provide suggestive evidence that the latter is likely the case; the increased likelihood of non-wage income was likely from reduced labor market attachment, increased probability of unemployment spells, and a greater likelihood of being on public work relief.

4.4 Robustness Checks

The full set of estimates indicates that compared to sons of non-fatal accident victims, sons who lost their fathers experienced reduced income during adulthood. The reduced income was driven largely by those who lost their fathers at an early age. In this section, I present some robustness checks meant to probe the paper's main results.

Potential Confounders

As described above, the empirical strategy relies on the assumption that the severity of an accident is unrelated to a father's characteristics that may also determine their son's future economic success. While Figure 3 shows the observed differences between fatal and non-fatal accident victims are small and statistically insignificant, one may still be concerned that certain unobservable traits could still jointly influence both accident severity and the long-term outcomes of the sons. A notable example of this concern is the potential influence of alcoholism, which could directly impact both the likelihood of experiencing a fatal accident

[20] While census enumerators did not collect detailed responses of non-wage income, they were instructed to ask if respondents had an income of more than $50 from sources other than their wages and salary. Specifically, enumerators counted profits, professional fees, collected rental income, interest, dividends, royalties, and payments-in-kind. Enumerators also classified unemployment compensation, government and private relief, pensions, and annuities as non-wage, non-salary income.

and the human capital formation of the victim's son. This may be particularly salient within the mining industry, where the rate of heavy alcohol consumption is more than double the national average.[21]

While I cannot directly test for differences in alcohol consumption between fatal and non-fatal accident victims, I am able to test for differences due to the availability of alcohol stemming from local prohibition policies. Specifically, I borrow from Howard and Ornaghi (2021) and account for the availability of alcohol with county-level prohibition policies enacted throughout the sample period.[22] To proxy for the local availability of alcohol at the time of each accident, I classify individual accidents as "dry" if they occur in a county after a local prohibition law was enacted. I then re-estimate the effect of losing a father on sons' wages using the preferred specification and include an indicator for whether their father's accident occurred during prohibition. Column 1 of Table A.7 shows that the estimates are nearly unchanged after accounting for variation in the local availability of alcohol; for instance, the average effect of losing a father after controlling for local prohibition is similar to those derived from the preferred specification.

Similarly, there is reason to be concerned that the cause of an individual accident may affect mortality and that selection into differing accident types may bias the estimates in ambiguous ways. Highlighted in Section 2, the likelihood of fatality varied depending on the cause of an accident. If, for example, falling victim to certain types of accidents was correlated with the skill of a miner, their work experience, or nativity, then there would be non-random sorting in the analysis sample. This concern is particularly relevant since miners often performed dissimilar tasks within their workplace, and these separate roles faced varied risks of accidents (Fishback, 1992). Put differently, miners involved in accidents caused by explosives might differ from those involved in accidents due to machinery failure, and such differences could influence the long-term outcomes of their sons. Consequently, the estimated effects of a father's death may capture this form of selection and bias the causal effect of losing a father.

To probe this potential source of selection bias, I first re-estimate the effect of a father's death using only data originating from Pennsylvania records, the only state to include complete accident causes. Column 2 of Table A.7 shows both the average effect of losing a father, and the effect of losing a father at a particularly young age on wages among descendants

[21] See, for example, Bush, D. and Lipari, R. (2015), *Substance Use and Substance Use Disorder By Industry*, Substance Abuse and Mental Health Services Administration. Accessed at: https://www.samhsa.gov/ (August 20, 2023)

[22] Before the Volstead Act in 1919, prohibition of alcohol was enacted locally, and between 1900 and 1919 nearly 60 percent of counties enacted a dry law—a law prohibiting the sale or consumption of alcohol. Howard and Ornaghi (2021) relies on local prohibition law changes collected by Sechrist (2012) to estimate the causal effects of Prohibition adoption on population sorting and farm values.

of victims identified in Pennsylvania; while the estimates are somewhat larger than the full sample, they are not statistically different. In Column 3, I include fixed effects for the cause of the accident and find that, even after accounting for differences by the type of accidents, the estimated effects are not meaningfully different from the Pennsylvania-only sample.

Finally, as mentioned above, the analysis includes childhood and adulthood county of residence fixed effects, which may raise separate concerns for estimation. For instance, a lingering concern is that childhood county fixed effects may fail to capture unique differences between mining towns. In other words, fathers identified within the same county may have differed in meaningful ways; the populations in smaller town mining operations in Allegheny County, PA, likely differed from those in Pittsburgh. Data from Berkes et al. (2023) instead allow for town-level fixed effects, and the fourth column of Table A.7 shows the main conclusions are not statistically different after replacing childhood county-level fixed effects with those at the childhood town level. Moreover, while the main specifications include adulthood county fixed effects to account for variation in income by geography, one may also be concerned that the adulthood county of residence reflects the sons' mobility and may be affected by losing a father. To address this, Column 5 omits adulthood county of residence from the regression and shows that the estimates are not meaningfully changed.

Outlier Ages

Another concern is that the estimated effects of parental death on a child's adulthood income could result from outliers among the relatively few sons who lost their fathers earlier during childhood. Within the analysis sample, roughly 1,725 sons lost their fathers due to a mining accident, and just 14 percent (N=245) of them experienced the death of a parent at an early age. I conduct a simple exercise and test whether one particular age group drives the estimated effect. Specifically, I re-estimate $\gamma + \theta$ from Eq. 2 and individually omit each age that comprises the definition of $Young_i$ at the time of the accident. Put differently, I re-estimate the effect of losing a parent during early childhood by leaving out sons who were infants when their father died, then by leaving out sons who were one year old when their father died, and so on.

Panel A of Figure A.3 shows the robustness of the effect of parental loss at an early age on income to outliers in age at the time of the accident. The leftmost estimate represents the baseline estimate and the other estimates present the results from omitting sons who were a specific age at the time of their father's accident. Panels B through D similarly report the results from these robustness exercises for the other measures of economic well-being. Altogether, the estimates are comparable in magnitude and are not significantly different, suggesting the effects of parental loss at an early age are not driven by outlier age groups in the definition of parental loss during early childhood.

Alternative Inference

As mentioned above, statistical inference is based on standard errors clustered at the adulthood county of residence level. Doing so allows for arbitrary correlation among those who reside in the same county during adulthood, such as local labor market shocks or economic conditions. However, one concern is that computing clustered standard errors at this level may ignore potentially important correlations across units. Failure to account for such correlation across units would result in biased estimates of standard errors and yield improper inference.

To address this concern, I probe the sensitivity of the main results to alternative levels of clustering and methods of inference. Specifically, I derive p-values from (i) clustering standard errors at the adulthood county of residence level from the primary analysis, (ii) clustering at the childhood county, (iii) two-way clustering by childhood and adulthood county, (iv) clustering by state-by-accident year cohort, and (v) wild-clustered bootstrapped standard errors. The first column of Table A.8 presents the estimates of the effect of parental loss during early childhood on adulthood wages and p-values derived from these alternative methods. Columns 2 through 4 present similar information for the other measures of economic well-being. The main conclusions remain unchanged regardless of the chosen inference method.[23]

4.5 Other Labor Market Outcomes

Reduced wage income is a proxy measure of decreased economic well-being and reflects bereaved sons' total labor market experience. For instance, bereaved sons may have held lower-earning occupations as adults or earned less within broadly similar occupations. In addition, reduced wage income may reflect lower labor market attachment or a greater likelihood of unemployment. This section estimates whether the observed differences in earnings are, at least in part, due to reduced labor market outcomes of the bereaved.

Occupation Sorting

As highlighted in Section 2, coal mining in the early 20^{th} century was the industry with the second highest rate of intergenerational transmission of occupation. As a result, the connection between father and son was likely an important link to the labor market. It is

[23] In Section C.4, I instead conduct inference by implementing a non-parametric placebo test of the effect of losing a father during early childhood on adulthood economic well-being and find similar conclusions.

then a natural question to consider whether or not fatal accidents disrupted the pathway into mining and affected occupational sorting. The first two columns of Table 5 show that compared to sons of non-fatal accident victims, those who lost their fathers were substantially less likely to work as miners as adults. In Panel A, the estimates suggest that, on average, those who lost their fathers were roughly 16 percent less likely to be miners or work in the industry themselves. The estimates presented in Panel B suggest a more significant shift away from the mines for those who were younger; those who lost their fathers during early childhood were nearly 50 percent less likely to work as miners.

To further examine how parental death affected occupation sorting, I group occupations into high-, semi-, and low-skill categories.[24] I then estimate the effect of losing a father on the probability of being employed in differently-skilled occupations. Columns 3 through 5 examine whether or not bereaved sons were differentially employed in high-, semi-, or low-skilled occupations as adults, respectively. Though Column 3 shows that those who lost their fathers did not sort into higher-skilled occupations, Column 4 shows that sons who lost their fathers were somewhat less likely to work in semi-skilled occupations. Though lacking in precision, the estimates are suggestive that sons who lost their fathers at an early age were five percentage points (roughly 10 percent) less likely to work in a semi-skilled occupation. Since this category includes mine operatives, the estimates likely reflect the reduced likelihood of becoming a miner. Column 5 shows that sons who lost their fathers perhaps entered low-skilled occupations. While the precision of the estimates does not allow for strong conclusions, they suggest that sons who lost their fathers at an early age were four and a half percentage points (20 percent) more likely to work in a low-skilled occupation.

The estimates of Table 6 are suggestive that sons who lost their fathers in a mining accident did not become miners themselves and were somewhat more likely to work in lower-skilled occupations. Consequently, the observed decline in wage income among those who lost their fathers may be, in part, due to employment in typically lower-earning occupations. To test this, I examine the sons' occupation-based income scores in the final column of the table. The estimates suggest that sons who lost their fathers experienced small differences in their income scores. At most, the point estimates suggest that sons who lost their fathers at an early age experienced a 2.7 percent decline in their occupational-based income score.

Why would sons who lost their fathers experience lower wage income and not a lower income score? First, the typical occupation-based income score tends to underestimate income from non-wage sources. If bereaved sons were more likely to enter occupations that derived

[24] High-skilled occupations comprise professionals, managers, and craftsmen. Semi-skilled occupations include sales workers, clerical workers, and operatives (including miners). Low-skilled occupations include farmers, service workers, and laborers. In Section C, I describe the aggregation of occupational codes in more detail.

a higher share of income from business or in-kind payments (e.g., farming or low-wage labor), then the occupation score may fail to capture these differences. A significant literature suggests alternative measures of occupation-based income scores to correct for differences in earnings by the source of income, industry, geography, and by race (Collins and Wanamaker, 2014; Saavedra and Twinam, 2020). In Table A.5, I instead implement suggested corrections to the occupation-based income score and show similar effects.

Given the estimated effects on income from non-wage sources, it is natural to wonder if the observed effects on income are driven by sorting into occupations known to earn a greater share of their incomes from non-wage sources like farming, for instance. Table A.6 examines the extent to which sons who lost their fathers were more likely to work in occupations known to earn a greater share of income from non-wage sources, general labor jobs, or farming. While there is some evidence that sons sorted into these occupations, it cannot explain the observed decline in wages.

The table shows that, on average, sons who lost their fathers were no different in their likelihood to work in occupations that received a greater share of income from non-wage sources, as laborers, or in farming. However, there is some evidence that sons who were young when their fathers were killed did. Those who lost their fathers earlier during childhood were nearly 20 percent more likely to work in a higher non-wage earning occupation and nearly 85 percent more likely to work as a farmer, though the estimates are not statistically significant. Still, sorting into these specific occupations is unlikely to explain the decline in wages among those who were young when their fathers died; in a separate regression controlling for working in higher non-wage earning occupations, sons who lost their fathers experienced similar losses in wage income (the point estimate is -0.163 with a corresponding standard error of 0.063). Likewise, even though the magnitude of the estimate of working as a farmer is large, the overall share in the sample is small, and dropping bereaved sons who were farmers as adults does not change the main conclusions of the paper.

Second, while wage income measures reported income across individuals, the income score measures the median income across occupations. Therefore, the income score effectively removes within-occupation income variation and only captures between-occupation differences in earnings. The fact that the estimated differences in income score are smaller in magnitude relative to the estimated wage effects suggests that sons who lost their fathers worked in professions with roughly similar median incomes but earned less within those occupations.

Employment

Bereaved sons may have experienced reduced wage income if the loss of their father is associated with reduced employment more broadly. For instance, sons who lost a father may have been less likely to actively participate in the labor force or have been more likely to be unemployed while in the labor force (Fronstin et al., 2001). To examine if differences in labor market attachment can be, at least partly, responsible for the estimated loss of income, I focus on three broad employment outcomes: labor force participation, unemployment, and employment on a public emergency project. I examine the association between losing a father and each employment measure in Table 6.

In the first column, I show that parental loss was not associated with any sizable change in labor force participation. Relative to an average labor force participation rate of 94 percent among sons of non-fatal accident victims, the point estimate of -0.8 percentage points is very close to zero. While sons who lost their fathers in a mining accident were not significantly more likely to be apart from the labor force, they were substantially more likely to report being out of work. Column 2 shows that sons who lost their fathers at an early age were about six percentage points (65 percent) more likely to report being unemployed at the time of the census.

Since orphaned sons were more likely to report being unemployed, they may also have been more likely to be employed by a public work relief program. Between 1933 and 1943, several New Deal agencies provided an hourly wage for work on public projects. These projects often involved labor-intensive work (e.g. construction of civil infrastructure), and payments were typically below wages available through private employment. Still, work relief was seen as an alternative to unemployment that provided a subsistence-level income to households (Fishback, 2017; Neumann et al., 2010). In the third column of Table 6, I regress an indicator for whether or not the respondent reported being employed by a public relief program on parental loss during childhood.[25] The estimates suggest that sons who lost their fathers during early childhood were also more likely to report being employed by a work relief program. Relative to sons of non-fatal accident victims, sons who experienced parental loss early during childhood were nearly three percentage points (20 percent) more likely to report employment in public emergency work. Given that relief work tended to be dominated by relatively low-wage, labor-oriented jobs and that the estimates of Table A.6 suggest that bereaved sons were not more likely to specifically work as general laborers, it is likely that those employed on relief work did so in a variety of occupations.

[25]There is considerable discussion regarding the measurement and classification of public work relief in the 1940 census. Given the uncertainty and error in recording public work relief, these estimates may be attenuated and should be viewed as a lower bound. I further discuss these issues of classifying work relief in Section C.3.

5 Channels

Altogether, the main estimates suggest bereaved sons experienced lowered economic well-being as adults. On average, sons who lost their fathers had roughly four-and-a-half percent lower wage income as adults compared to those whose fathers survived. These estimates contain important heterogeneity; there is some suggestive evidence that the availability of alternative caretakers present in the childhood home mitigated the long-run effects of losing a father, and sons whose fathers died earlier during childhood experienced substantial long-run effects, having lower wage income on the order of roughly 16 percent. However, understanding the precise intermediate channels that led sons who lost their fathers, and in particular, those who did so earlier during childhood, to experience broadly reduced economic well-being is difficult.

The extensive literature on child development, family shocks, and long-run outcomes suggests some key pathways. Importantly, the literature on childhood development and human capital formation unambiguously highlights the important role of broad parental investments. Specifically, the time and care provided by parents are arguably the most valuable input to the development of children, influencing educational attainment, cognitive and non-cognitive skills, and welfare later in life (Cunha and Heckman, 2008; Cunha et al., 2010; Heckman et al., 2013; Seror, 2022). Beyond parental time, parents' monetary and material investments also play a key role in enhancing a child's human capital and long-term well-being (Dahl and Lochner, 2012; Del Boca et al., 2014). Additionally, the marginal effects of parental investment appear dependent on a child's age, suggesting critical stages of child development (Cunha et al., 2010; Currie and Almond, 2011; Del Boca et al., 2014; Heckman et al., 2013; Heckman and Mosso, 2014). While their dynamic complementarities complicate the relative importance of parental inputs, some research suggests parental time, especially earlier during childhood, may be more productive in shaping child development (Del Boca et al., 2014).

That the largest effects are observed for those who lost their fathers earlier during childhood suggests disruptions to childhood development and the family environment play a key role. Similar to the broad literature on child development, the literature on early childhood shocks argues disruptions (and improvements) during critical ages of child development can have wide effects (Currie and Almond, 2011). For instance, sharp changes to the home environment, a parent's employment, and access to healthcare have been shown to affect a child's human capital accumulation, physical health, and adulthood labor market outcomes (Andrabi et al., 2021; Currie and Almond, 2011; Grönqvist et al., 2020; Hoehn-Velasco, 2021; Ruhm, 2004). Research has similarly identified a range of outcomes affected by the death of a parent, which are likely to matter for long-run well-being. The loss of a parent, especially in the context of developing countries, is associated with disruptions to human capital formation, reducing educational attainment and quality, and leading to children forgoing school for income-earning occupations (Case and Ardington, 2006; Chen et al., 2009; Gertler et al., 2004; Gimenez et al., 2013). Moreover, losing a parent comes with serious psychic costs for surviving family members that likely augment cognitive and behavioral development during important stages of a child's development (Böckerman et al., 2023; Guldin et al., 2015; Høeg et al., 2018).

In the context of losing a father due to a mining accident during the early 20$^{\text{th}}$ century, there are potentially several pathways through which bereavement affects the long-run economic outcomes of children. For instance, given the gendered division of labor at the time, losing a father resulted in a significant shock to household resources. Beyond the loss of income, losing a husband required a reallocation of time for the surviving widows and children, potentially shifting away from domestic production or schooling toward the labor force (McGill, 1923). Losing a husband also likely resulted in eventual remarriage and a change in household composition. Moreover, since the mining industry saw high rates of intergenerational transmission of occupation, the loss of a father may have been particularly salient. For bereaved sons, losing a father likely disrupted a key channel of skill accumulation, occupation-specific knowledge, and labor-market connections.

While the historical data and context of mining accidents allow for careful identification of the long-run net effects of losing a father, tracing out the intermediate channels responsible for the observed causal effects in this setting is a significant challenge due to the paucity of outcomes. Historical census data do not permit a detailed investigation of some potentially important pathways highlighted by the literature. For instance, census data contains no metric to proxy for the psychic costs of losing a loved one, no information about parental time or material resources allocated throughout childhood, and no information about cognitive or non-cognitive skills.

Still, one can use census data to observe some proxies for likely important intermediate channels following the death of a parent. As described in Section 3, I rely on a sample of sons of accident victims identified in the nearest census following their father's accident, as well as mothers whose husbands were involved in a mining accident. Doing so allows me to examine broad proxies for parental investment, differences in schooling attendance and labor force participation later during childhood, and educational attainment.

5.1 Mothers' Outcomes

Coal families followed a gendered division of labor, where husbands worked to earn wages and domestic production belonged to the wives and children (Giesen, 1995; Roberts, 1904; Wilson, 1922). Consequently, the death of a husband meant a substantial shock to household resources. Although the families of accident victims often received compensation to defray some of the lost income, benefits were just a fraction of typical earnings and unavailable for the long term. As a result, the capacity for broad parental investments made by the surviving widow, both time and material, likely shifted substantially.

First, I show that women who lost their husbands were four and a half times more likely to report being widowed during the nearest decennial census. Specifically, while just five percent of wives of non-fatal accident victims reported being widowed, I show in Column 1 of Table 7 that women who lost their husbands were nearly 24 percentage points more likely to be widowed. Next, Column 2 shows that roughly 30 percent of women who lost their husbands in a mine accident reported being the head of household in the nearest census following the accident. Compared to wives of non-fatal accident victims, bereaved mothers were 25 percentage points more likely to be the head of the household.

One way to view the significant share of women who lost their husbands reporting as a widowed head of household is as an alternative to remarriage. Put differently, compared to wives whose husbands survived, those women whose husbands were killed were substantially less likely to be married. Though roughly one-quarter of these women had not yet remarried, the remainder did, implying a change to the family environment following a miner's death. On the one hand, widows who did not marry may have had fewer children, meaning the surviving bereaved children had fewer siblings to share scarce parental time and household resources. Alternatively, remarriage may have resulted in joined households, increasing the number of children in the home. However, Column 3 examines differences in the number of children in the household and shows that mothers who lost their husbands had similar family sizes following their husbands' accidents.

Finally, in Column 4, I examine whether mothers who lost their husbands in an accident entered the labor force. The estimates show that women who lost their husbands were about four percentage points more likely to report being in the labor force. Compared to women whose husbands survived, these estimates suggest that the death of a child's father is associated with a 63 percent increase in the likelihood of labor force participation. Most of those who entered the labor force worked in services for private households (e.g., launderers and cleaners). While these women were more likely to enter the labor force, it is unlikely doing so replaced lost income.

The set of estimates implies that the surviving households shifted significantly following the fatal accident. As widowed mothers had to cope with the loss of their husbands, they were substantially more likely to enter the labor force in favor of an income-earning occupation and remained the single head of their household more than the counterfactual. As a result, the death of a husband likely resulted in both income (loss of household resources) and substitution effects (loss of parental time) for investment in the bereaved children. Ultimately, the historical data do not allow one to disentangle the two; the census records contain no metric of time spent caring for children in the home nor full compensation received by widows.[26] Still, there is at least some suggestive evidence that changes in parental time played an important role. While the estimate's precision does not allow for strong conclusions, the estimates in Table 3 show that sons who lived in households with extended family members when they lost their fathers experienced no differences in wages as adults. Given that the broad literature on child development highlights the outsized role of parent-child interactions and parental time, and the gender-specific effects of losing a father, it is possible that changes in parental investments were most strongly affected by the loss of the time a son spent with their father (Del Boca et al., 2014; Kalil et al., 2016; Weinberg et al., 2019).

Finally, the fact that mothers are identified by living in the same household with their sons in the nearest census following an accident raises concerns about selection bias. For instance, losing a husband may affect families' living arrangements, resulting in children living in households apart from their mothers for unobserved reasons. Alternatively, losing a father may induce children to remain in the household to support the family when they would have otherwise departed. While it is difficult to quantify the degree to which this sort of selection may affect the overall estimates, in Table A.9, I examine a subset of mothers identified by children who were still young enough in the nearest census following the accident where living apart would be unlikely. Specifically, I re-estimate the outcomes of mothers identified via sons who were 18 or younger in the census following their fathers' accidents and find similar effects.[27]

The results of this section show that those widowed by a mining accident took on an increased burden. These mothers were substantially more likely to join the labor force and be listed as the head of their household during an era where single-mother households were largely uncommon. While the context makes it difficult to say whether losing a father reduced

[26] Moreover, disentangling the loss of income of a deceased parent from the reallocation of parental time requires additional variation, for instance, in the household finances following the death of a parent. In other words, it is not enough to have quasi-random variation in the mortality of a parent; one needs an additional arm of variation in either time spent with children or household resources.

[27] I also directly estimate the probability that the son is observed in the same household as their mother in the census following an accident. I find a point estimate of -0.018 with a corresponding standard error of 0.008, suggesting that bereaved sons are less likely to be co-resident with their mothers following an accident. Though the control group mean is 85 percent, so the percent effect is of very small magnitude.

parental investment through household resources or parental time, there is broad evidence both likely shifted. Taken together, the observed declines in adulthood economic well-being from parental loss may reflect, at least in part, a reduced capacity for parental investment on the part of the surviving mother.

5.2 Childhood Outcomes

If the surviving family required bereaved sons to reduce schooling to supplement household resources, the observed decline in adulthood wage income might reflect broad disruptions to human capital formation. On the other hand, an early entry into the labor market may have been beneficial if it provided young men with early experience and career opportunities. In this case, the estimated null effect of parental loss among those relatively older at the time of their father's accident could reflect forgone returns to schooling in favor of additional work experience. To examine disruptions to human capital accumulation during childhood, I focus on the sons of accident victims followed to the nearest census after their father's accident as described in Section 3.

The first column of Table 8 shows differences in labor force participation of sons after their father's accident. Among those who did not lose their father in an accident, roughly 35 percent were in the labor force during the nearest census following the accident. Sons who lost their fathers during childhood reported no difference in their supply of labor following the accident. While the estimated coefficient implies that sons who lost their fathers in an accident were one percentage point more likely to report being in the labor force, the estimate is small in magnitude (roughly three percent) and not statistically significant.

While sons who lost their fathers may not have differed in their likelihood of being in the labor force following an accident, they may have been more likely to work in the mine, reflecting a need for immediate substitution of their father's labor. Column 2 shows this was not the case; there is no consistent, large, or significant evidence that those who lost their fathers were more likely to work in mining soon after their father's accident. Alternatively, sons who lost their fathers may have experienced a disruption to schooling later on during childhood and following their father's accident. In Column 3, I examine differences in whether or not sons of accident victims reported attending school in the nearest census after the accident. The estimates suggest that those who lost their fathers did not differ in their likelihood of attending school following an accident.

Next, I examine a broad proxy for human capital formation: milestones of school completion and total educational attainment observed during adulthood. In Columns 4 and 5, I show that nearly 80 and 25 percent of sons in the control group reported completing the 8$^{\text{th}}$

grade and high school, respectively. Estimates show that sons who lost their fathers were not significantly different in their completion of both schooling milestones. Finally, in Column 6, I estimate differences in total years of schooling between sons who lost their fathers in a mining accident and those whose fathers survived. The column shows that bereaved sons were not meaningfully different in their total years of education. Why might educational attainment not differ among those who lost their fathers at an early age? Benefits to widows were implicitly meant to keep children in school. For instance, Roberts (1904) highlights mine operators claimed, in the case of a fatal accident, "If there are children, provision is made for their education." Furthermore, the sample period coincides with the rise of, and expansions to, state mandates for compulsory education, which assured that children remained in school at least until age fourteen.

Taken together, there is little to suggest that bereaved children experienced large disruptions to their human capital accumulation in this context. Although the available census data allows for the estimation of differences in total years of education, it falls short in capturing aspects crucial to human capital development, such as the quality of education, which are suggested by the bereavement literature to be impacted by the loss of a parent (Case and Ardington, 2006; Chen et al., 2009; Gimenez et al., 2013). Additionally, given the high rates of intergenerational occupation transmission among coal miners, fathers likely provided occupation-specific knowledge, and losing a father disrupted such a chain of tacit knowledge. Ultimately, the data are not rich enough to estimate specific measures of human capital that are likely to inform the long-run wage effects.

6 Conclusion

Motivated by its importance, a large body of research has sought to document the consequences of disruptions to family structure on a child's adulthood outcomes (Amato and Anthony, 2014; Bloome, 2017; Corak, 2001; Gruber, 2004; Kalil et al., 2016; McLanahan et al., 2013; Weinberg et al., 2019). Yet the long-run effects of the most severe disruption to a child's family—the death of a parent—are not well understood due to data limitations and the difficulty of identifying a valid counterfactual for bereaved children.

This study addresses this gap in the literature and provides an estimate of the long-run, causal effects of losing a parent. Specifically, it leverages individual-level, linked U.S. Census data along with a novel identification strategy to estimate the effects of the loss of a father on the adulthood economic well-being of bereaved sons. I identify the effects of parental loss by using mining accident records from the early 20^{th} century that allow me to compare the adulthood outcomes of sons who lost their fathers in a mining accident to sons whose fathers survived. Importantly, the constructed dataset of linked census micro-data allows me to follow sons of accident victims to the 1940 census, which directly captures wage income, educational attainment, and broad labor market outcomes.

The estimates presented in this paper suggest that the loss of a father has far-reaching effects and reduces the bereaved son's adulthood economic well-being. Sons who lost their fathers experienced a 4.5 percent decline in annual wage income—roughly 1.25 years of lost income—as adults. This average effect masks important heterogeneity; sons who were particularly young when they lost their fathers experienced an income loss of 16.5 percent. Moreover, while it is difficult to identify clear-cut pathways of occupational sorting, there is some evidence that bereaved sons experienced other meaningful differences in the labor market; those who lost their fathers were less likely to become miners themselves and were more likely to report spells of unemployment and be employed on relief work.

While disentangling the intermediate channels through which losing a father likely mattered for long-run economic well-being is challenging, I test several key outcomes highlighted by the extensive literature on child development and human capital formation. Specifically, I show that orphaned sons grew up in households with a broadly reduced capacity for parental investment. Wives of miners who perished in an accident were substantially more likely to join the labor force and be listed as the single head of household. Given the strict, gendered division of household production among coal families of the era, the surviving widows more often had to play the role of income earning and child raising on their own, and in an era of considerably less public assistance. However, there is little evidence that sons who lost their fathers experienced disrupted human capital attainment. Bereaved sons were equally likely to be attending school when observed following the accident, were not more likely to enter the labor force early, and, ultimately, completed similar years of education.

Still, the precise intermediate mechanisms are difficult to determine in this setting. While studying historical mine accidents allows for a novel design to uncover the long-run effects of parental loss, it does not allow an in-depth examination of the mechanisms behind its net effects. Education quality (Gimenez et al., 2013), parent-specific human capital (Kalil et al., 2016), and mental health (Berg et al., 2016; Böckerman et al., 2023; Guldin et al., 2015; Høeg et al., 2018) have been shown to both play an important role in human capital formation and be affected by the death of a parent. However, the historical setting does not permit an examination of these channels. Though the estimates suggest that, following an accident, bereaved sons are in school at equal rates to those whose fathers survived an accident, the data cannot speak to education quality. Similarly, the data contain no proxy measure of behavioral well-being or parental time, and ultimately, the context does not allow a full examination of cognitive, non-cognitive, or behavioral skills that may be particularly affected by losing a father.

Examining parental death in the early 20th century with linked census micro-data and death from mining accidents also limits the analysis to focus on fathers and sons. Of course, the causal effects of parental death may differ for daughters. For instance, Gimenez et al. (2013) shows parental loss, particularly a father's death, affects daughters' educational attainment in the short-run more strongly than sons and may increase the probability of marrying young. Additionally, Cas et al. (2014) suggests that parental death may cause daughters to take on a supporting role in the household, helping to care for the surviving family at the cost of school enrollment. These effects may be particularly relevant in the context of mining families where household responsibilities were shared after the death of a father, and daughters may have had to play an increased role in the household (Lantz and McCrary, 1958).

Though the historical context brings several limitations, coal families in the early 20th do provide a unique setting to highlight the importance of parental relationships in shaping the later-in-life well-being of children. The relationship between fathers and sons was likely particularly important among coal families. At the time, coal mining was an industry with particularly high rates of intergenerational transmission of occupation; sons of miners entered their fathers' occupation at nearly double the rate of other families. Additionally, the setting is more similar to those with fewer mechanisms for public financial support. Estimating the causal effects of parental loss in the modern context is challenging since there exists generous support for orphaned children, and estimates of parental loss cannot separate resources meant to alleviate the effects of losing a parent (Adda et al., 2011; Böckerman et al., 2023). In contrast, due to the lack of formal social support systems for families of accident victims, the estimates in this study are more likely to represent the long-run effects of the death of a father in the absence of a strong social safety net.

References

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2021). Automated linking of historical data. *Journal of Economic Literature 59*(3), 865–918.

Adams, W. W., L. Chenoweth, and L. E. Geyer (1932). *Coal-mine accidents in the United States; 1930.* U.S. Gov. Printing Office.

Adda, J., A. Björklund, and H. Holmlund (2011). The Role of Mothers and Fathers in Providing Skills: Evidence from Parental Deaths. *IZA Discussion Papers* (5425).

Amato, P. R. (2000). The consequences of divorce for adults and children. *Journal of marriage and family 62*(4), 1269–1287.

Amato, P. R. and C. J. Anthony (2014). Estimating the effects of parental divorce and death with fixed effects models. *Journal of Marriage and Family 76*(2), 370–386.

Andrabi, T., B. Daniels, and J. Das (2021). Human capital accumulation and disasters: Evidence from the pakistan earthquake of 2005. *Journal of Human Resources*, 0520–10887R1.

Bailey, M., C. Cole, and C. Massey (2020). Simple strategies for improving inference with linked data: A case study of the 1850–1930 ipums linked representative historical samples. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 53*(2), 80–93.

Bailey, M. J., C. Cole, M. Henderson, and C. Massey (2020, December). How well do automated linking methods perform? lessons from us historical data. *Journal of Economic Literature 58*(4), 997–1044.

Bailey, M. J. and W. J. Collins (2006). The wage gains of african-american women in the 1940s. *The Journal of Economic History 66*(3), 737–777.

Beegle, K., J. De Weerdt, and S. Dercon (2010). Orphanhood and human capital destruction: Is there persistence into adulthood? *Demography 47*(1), 163–180.

Berg, L., M. Rostila, and A. Hjern (2016). Parental death during childhood and depression in young adults–a national cohort study. *Journal of child psychology and psychiatry 57*(9), 1092–1098.

Berkes, E., E. Karger, and P. Nencka (2023). The census place project: A method for geolocating unstructured place names. *Explorations in Economic History 87*, 101477. Methodological Advances in the Extraction and Analysis of Historical Data.

Bloome, D. (2017). Childhood Family Structure and Intergenerational Income Mobility in the United States. *Demography 54*(2), 541–569.

Böckerman, P., M. Haapanen, and C. Jepsen (2023). Dark passage: Mental health consequences of parental death. *American Journal of Health Economics*.

Brophy, J. (1964). *A Miner's Life*. University of Wisconsin Press.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Cas, A. G., E. Frankenberg, W. Suriastini, and D. Thomas (2014). The impact of parental death on child well-being: evidence from the indian ocean tsunami. *Demography 51*(2), 437–457.

Case, A. and C. Ardington (2006). The impact of parental death on school outcomes: Longitudinal evidence from south africa. *Demography 43*(3), 401–420.

Case, A., C. Paxson, and J. Ableidinger (2004). Orphans in africa: Parental death, poverty, and school enrollment. *Demography 41*(3), 483–508.

Census Bureau (1940). Instructions to enumerators, population and agriculture, 1940. *1940 Census Publications Form PA-1*.

Census Bureau (1946). *Population: Internal Migration 1935 to 1940*. United States Government Printing Office.

Chen, S. H., Y.-C. Chen, and J.-T. Liu (2009). The impact of unexpected maternal death on education: First evidence from three national administrative data links. *American Economic Review: Papers & Proceedings 99*(2), 149–53.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American economic review 99*(4), 1145–77.

Collins, W. J. and M. H. Wanamaker (2014). Selection and economic gains in the great migration of african americans: new evidence from linked census data. *American Economic Journal: Applied Economics 6*(1), 220–252.

Conyngton, M. K. (1917). *Effect of Workmen's Compensation Laws in Diminishing the Necessity of Industrial Employment of Women and Children: Bulletin of the United States Bureau of Labor Statistics, No. 217*. U.S. Gov. Printing Office.

Corak, M. (2001). Death and Divorce: The Long-Term Consequences of Parental Loss on Adolescents. *Journal of Labor Economics 19*(3), 682–715.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Currie, J. and D. Almond (2011). Human capital development before age five. In *Handbook of labor economics*, Volume 4, pp. 1315–1486. Elsevier.

Currie, J. and C. Spatz Widom (2010). Long-term consequences of child abuse and neglect on adult economic well-being. *Child maltreatment 15*(2), 111–120.

Dahl, G. B. and L. Lochner (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review 102*(5), 1927–56.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

Dix, K. (1978). *Work Relations in the Coal Industry: The Hand-Loading Era, 1880-1930.* Number 7-2 in 78. West Virginia University Bulletin.

Dribe, M., E. Debiasi, and B. Eriksson (2022). The effect of parental loss on social mobility in early twentieth-century sweden. *Demography 59*(3), 1093–1115.

Dupraz, Y. and A. Ferrara (2021). Fatherless: The long-term effects of losing a father in the us civil war. *Competitive Advantage in the Global Economy (CAGE) Working Paper Series* (538).

Ferman, B. and C. Pinto (2019). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Review of Economics and Statistics 101*(3), 452–467.

Ferrie, J. P. (1996). A new sample of males linked from the public use microdata sample of the 1850 us federal census of population to the 1860 us federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 29*(4), 141–156.

Fishback, P. (2017). How successful was the new deal? the microeconomic impact of new deal spending and lending policies in the 1930s. *Journal of Economic Literature 55*(4), 1435–85.

Fishback, P. V. (1986). Workplace safety during the progressive era: Fatal accidents in bituminous coal mining, 1912–1923. *Explorations in Economic History 23*(3), 269–298.

Fishback, P. V. (1992). *Soft Coal, Hard Choices: The Economic Welfare of Bituminous Coal Miners, 1890-1930*. Oxford University Press.

Fishback, P. V. and S. E. Kantor (1998). The adoption of workers' compensation in the united states, 1900–1930. *The Journal of Law and Economics 41*(2), 305–342.

Fishback, P. V. and S. E. Kantor (2000). *A prelude to the welfare state: The origins of workers' compensation*. University of Chicago Press.

Fronstin, P., D. H. Greenberg, and P. K. Robins (2001). Parental disruption and the labour market performance of children when they reach adulthood. *Journal of Population Economics 14*(1), 137–172.

Gertler, P., D. I. Levine, and M. Ames (2004). Schooling and parental death. *Review of Economics and Statistics 86*(1), 211–225.

Giesen, C. A. B. (1995). *Coal Miners' Wives*. The University Press of Kentucky.

Gimenez, L., S.-Y. Chou, J.-T. Liu, and J.-L. Liu (2013). Parental loss and children's well-being. *Journal of Human Resources 48*(4), 1035–1071.

Grönqvist, H., J. P. Nilsson, and P.-O. Robling (2020). Understanding how low levels of early lead exposure affect children's life trajectories. *Journal of Political Economy 128*(9), 3376–3433.

Gruber, J. (2004). Is making divorce easier bad for children? the long-run implications of unilateral divorce. *Journal of Labor Economics 22*(4), 799–833.

Guldin, M.-B., J. Li, H. S. Pedersen, C. Obel, E. Agerbo, M. Gissler, S. Cnattingius, J. Olsen, and M. Vestergaard (2015). Incidence of suicide among persons who had a parent who died during their childhood: a population-based cohort study. *JAMA Psychiatry 72*(12), 1227–1234.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–86.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

Høeg, B. L., C. Johansen, J. Christensen, K. Frederiksen, S. O. Dalton, A. Dyregrov, P. Bøge, A. Dencker, and P. E. Bidstrup (2018). Early parental loss and intimate relationships in adulthood: A nationwide study. *Developmental Psychology 54*(5), 963.

Hoehn-Velasco, L. (2021). The long-term impact of preventative public health programs. *The Economic Journal 131*(634), 797–826.

Howard, G. and A. Ornaghi (2021). Closing time: The local equilibrium effects of prohibition. *The Journal of Economic History 81*(3), 792–830.

Inwood, K., C. Minns, and F. Summerfield (2019). Occupational income scores and immigrant assimilation. evidence from the canadian census. *Explorations in Economic History 72*, 114–122.

Kalil, A., M. Mogstad, M. Rege, and M. E. Votruba (2016). Father Presence and the Intergenerational Transmission of Educational Attainment. *Journal of Human Resources 51*(4), 869–899.

Lang, K. and J. L. Zagorsky (2001). Does growing up with a parent absent really hurt? *The Journal of Human Resources 36*(2), 253–273.

Lantz, H. and J. McCrary (1958). *People of Coal Town*. Columbia University Press.

MacKinnon, J. G. and M. D. Webb (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics 32*(2), 233–254.

McGill, N. P. (1923). *The Welfare of Children in Bituminous Coal-Mining Communities in West Virginia. United States Children's Bureau*. Number 117. US Government Printing Office.

McLanahan, S., L. Tach, and D. Schneider (2013). The Causal Effects of Father Absence. *Annual Review of Sociology 39*, 399–427.

Neumann, T. C., P. V. Fishback, and S. Kantor (2010). The dynamics of relief spending and the private urban labor market during the new deal. *The Journal of Economic History 70*(1), 195–220.

Roberts, P. (1904). *Anthracite Coal Communities*. The Macmillan Company.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021). IPUMS USA: Version 11.0 [dataset]. *Minneapolis, MN: IPUMS, https://doi.org/10.18128/D010.V11.0*.

Ruhm, C. J. (2004). Parental employment and child cognitive development. *Journal of Human resources 39*(1), 155–192.

Saavedra, M. and T. Twinam (2020). A machine learning approach to improving occupational income scores. *Explorations in Economic History 75*, 101304.

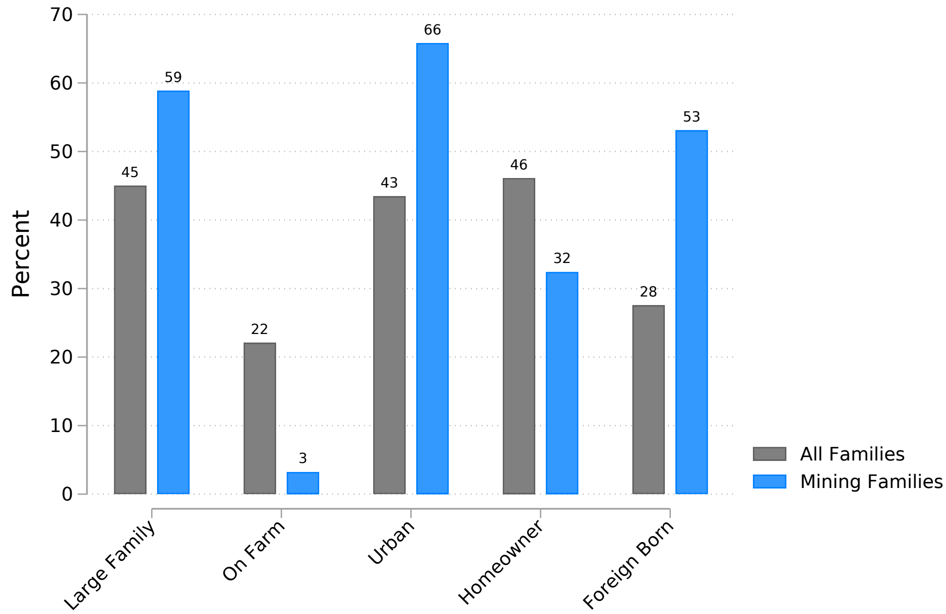
Sechrist, R. P. (2012). Prohibition Movement in the United States, 1801-1920.

Seror, A. (2022). Child development in parent-child interactions. *Journal of Political Economy 130*(9), 000–000.

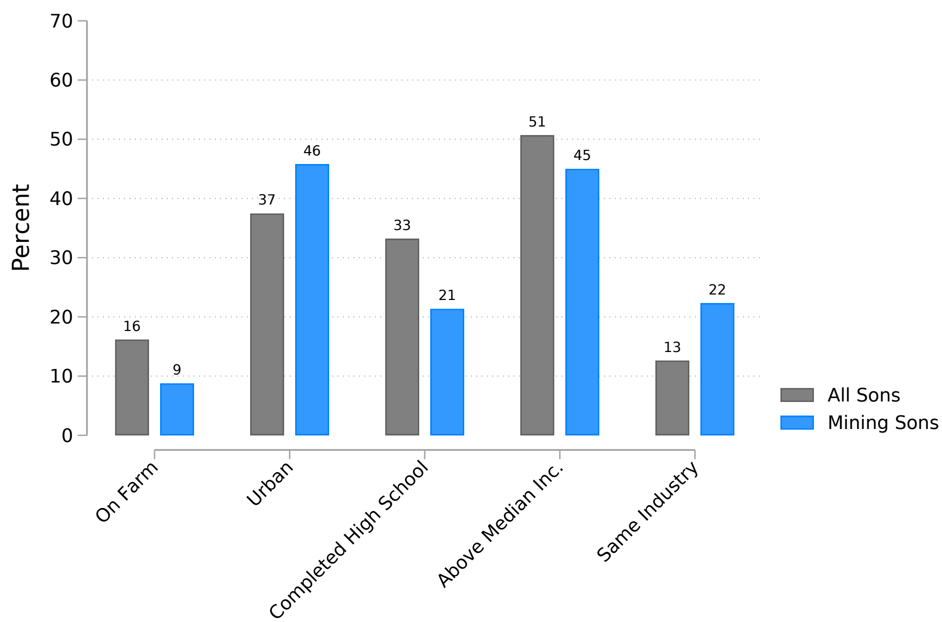
U.S. Commissioner of Labor (1908). *Twenty-Third Annual Report. Workmen's Insurance and Benefit Funds in the United States.* . United States Government Printing Office.

Weinberg, B. A., E. D. Gould, and A. Simhon (2019). Does Parental Quality Matter? Evidence on the Transmission of Human Capital Using Variation in Parental Influence from Death, Divorce, and Family Size. *Journal of Labor Economics*.

Wilson, H. (1922). *Child Labor and the Welfare of Children in an Anthracite Coal-Mining District. United States Children's Bureau.* Number 106. US Government Printing Office.



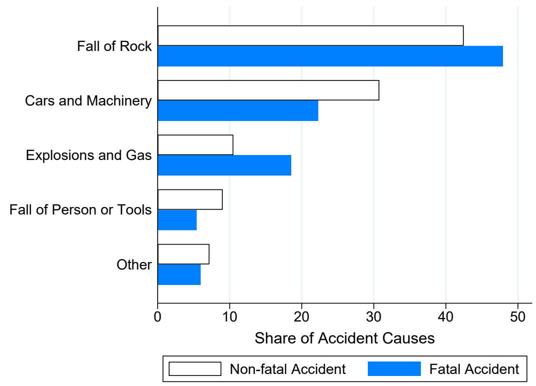
(a) Coal Mining Families



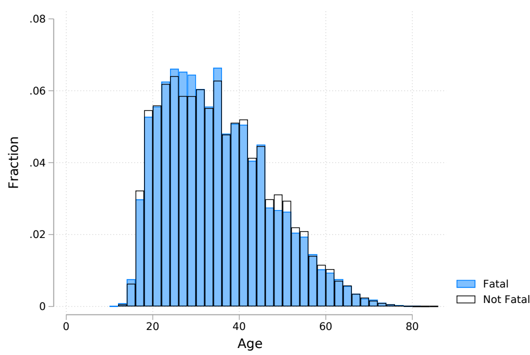
(b) Sons of Coal Miners in 1940

Figure 1: Coal Mining Families in the Early 20$^{\text{th}}$ Century

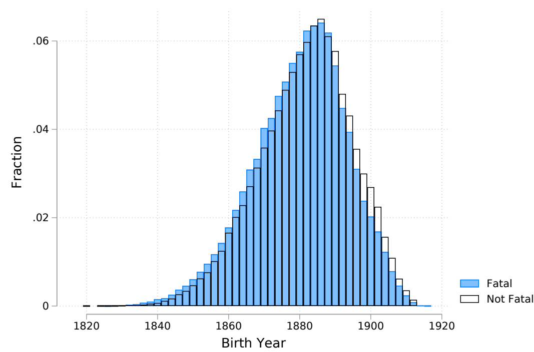
Notes: The first panel of the figure compares the average characteristics of household heads of coal mining families to those working in other industries observed in the 1900-1920 decennial censuses in the four core states. The second panel similarly compares the adulthood characteristics observed in the 1940 census of sons from coal mining families to those from families working in other industries. Large families are defined as a household head reporting having three or more children (the median family had two children in their household). Sons are recorded as working in the same industry as their fathers if they worked in the same 1950 Census Bureau industrial classification code.



(a) Accident Causes



(b) Histogram of Victim's Age



(c) Histogram of Victim's Birth Year

Figure 2: Accident Records by Fatality

Notes: The figure shows the nature and causes of fatal and non-fatal mine accidents as well as the distributions of victim's age and birth year listed in the accident records. The first panel describes the five main accident causes. Panels B and C of the figure present the distributions of age at the time of accident and birth year as listed in the accident records. The distributions of fatal accident victims are presented by filled blue bars while the distributions of non-fatal accident victims are presented by hollow bars.

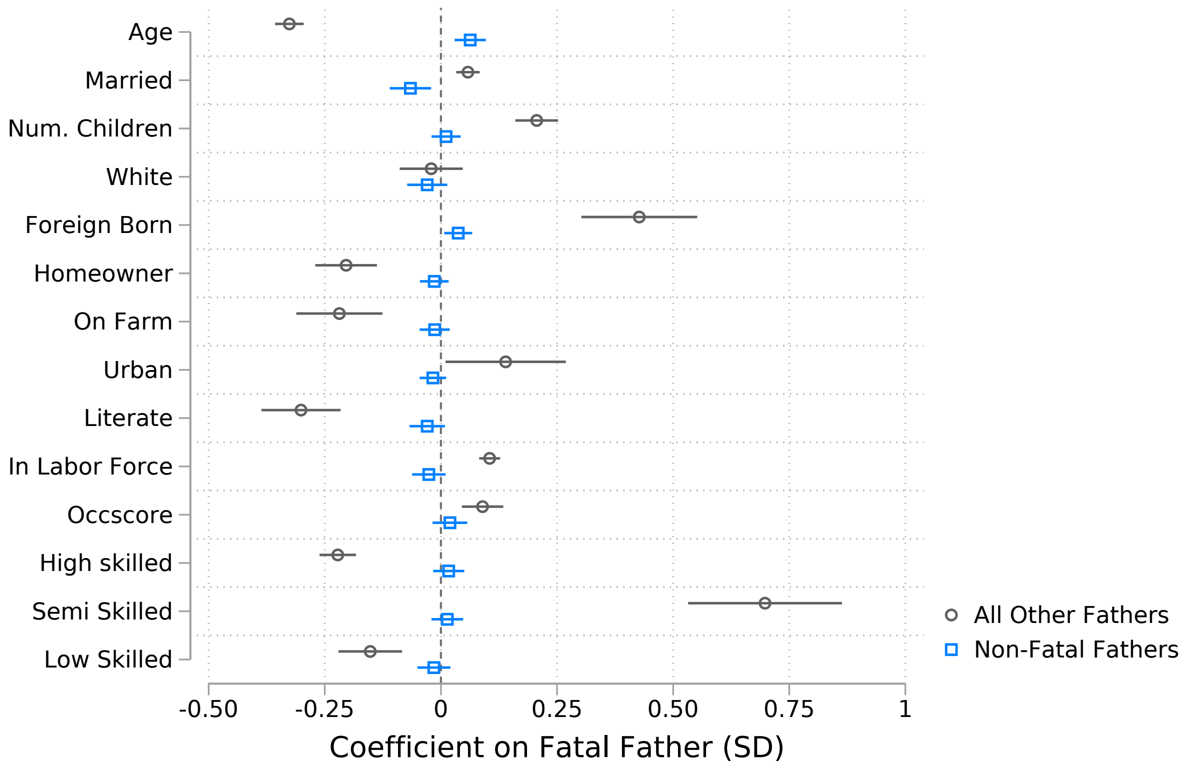


Figure 3: Difference in Means, SD of Control Group

Notes: The figure shows comparisons between fathers that eventually perish in a fatal mining accident and i) all other fathers residing in Pennsylvania, Illinois, West Virginia and Ohio and ii) fathers that are eventually involved in serious, but non-fatal mining accidents. I standardize the characteristics by the mean and standard deviation of each comparison group and plot the standardized coefficients (along with 90 percent confidence intervals) from a regression of each characteristic on a dummy for fatal accident victims conditional on census year and state of residence.

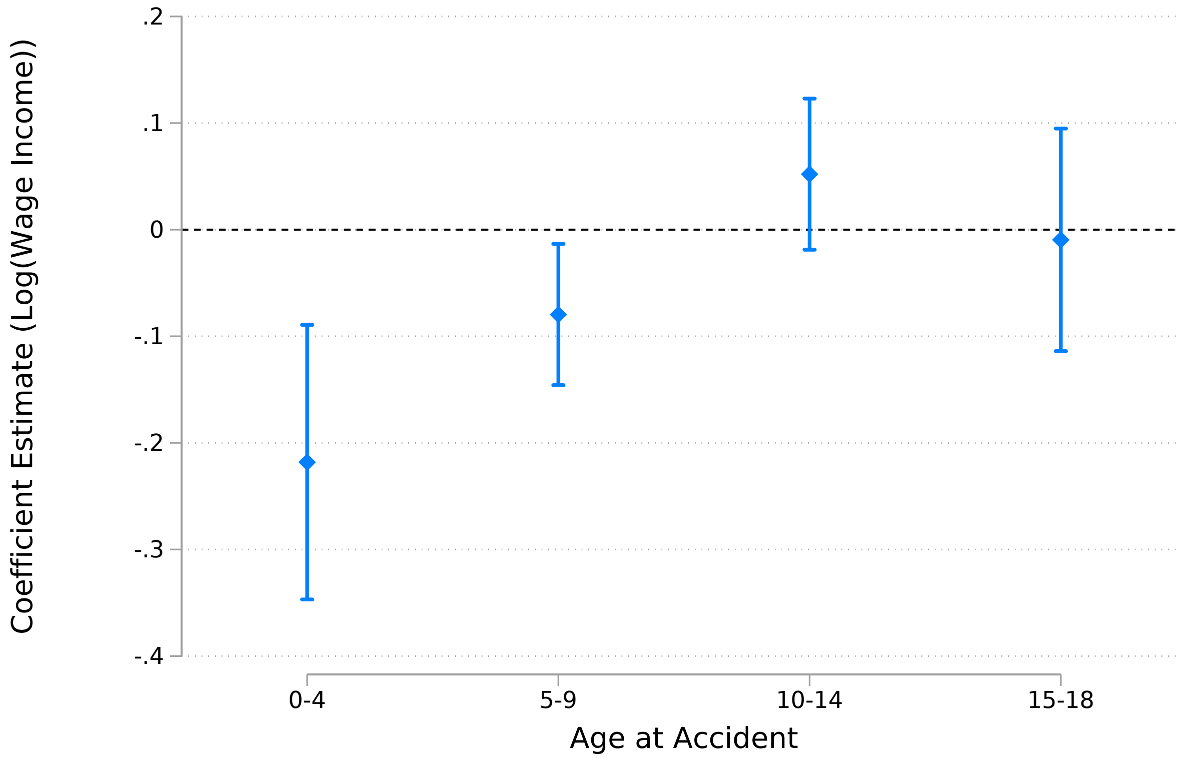


Figure 4: Parental Death on Wages by Age At Accident

Notes: The figure presents estimates of the average effect of parental death on sons' wage income by age at the time of the accident. Point estimates are shown as solid points, and 90 percent confidence intervals are plotted as lines.

Table 1: Summary Statistics of Accident Victims

	(1) Non-Fatal Fathers	(2) Fatal Fathers	(3) Differences
Age	38.44	38.64	0.617***
	[9.76]	[9.92]	(0.167)
Married	0.97	0.97	-0.010***
	[0.16]	[0.18]	(0.003)
Number of Children	3.17	3.22	0.022
	[1.96]	[1.95]	(0.033)
White	0.98	0.98	-0.004
	[0.12]	[0.15]	(0.002)
Foreign Born	0.47	0.48	0.019*
	[0.50]	[0.50]	(0.008)
Homeowner	0.37	0.34	-0.007
	[0.48]	[0.47]	(0.008)
On Farm	0.10	0.10	-0.004
	[0.29]	[0.29]	(0.005)
Urban Area	0.50	0.52	-0.009
	[0.50]	[0.50]	(0.008)
Literate	0.88	0.86	-0.009
	[0.32]	[0.34]	(0.006)
In Labor Force	0.98	0.98	-0.003
	[0.12]	[0.14]	(0.002)
Occupation Score	21.07	21.34	0.206
	[10.56]	[10.51]	(0.179)
High Skill	0.15	0.17	0.006
	[0.36]	[0.37]	(0.006)
Semi Skill	0.50	0.49	0.007
	[0.50]	[0.50]	(0.008)
Low Skill	0.26	0.25	-0.007
	[0.44]	[0.43]	(0.007)
Miner (Occupation)	0.42	0.41	0.006
	[0.49]	[0.49]	(0.008)
N	14,968	4,804	19,772

Notes: The table compares the mean characteristics (and standard deviation in brackets) of fathers listed in mining accident records in Pennsylvania, Illinois, West Virginia, and Ohio during the 1900, 1910, and 1920 censuses. The characteristics of victims who survived their accident are shown in the first column, while Column 2 similarly shows the characteristics of fathers who perished in a mining accident. The third column displays the differences in means between the two groups, conditional on census year and state of residence fixed effects. Standard errors are clustered by county of residence and are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Estimates of Losing a Father on Wage Income

	(1)	(2)	(3)	(4)	(5)
	Log(Wage Income)				
Father Fatal Accident	-0.026	-0.041*	-0.037	-0.037	-0.046*
	(0.024)	(0.023)	(0.024)	(0.023)	(0.026)
Control Group Mean	1,109	1,109	1,109	1,109	1,109
N	5,624	5,591	5,536	5,528	5,526
Census Year and County FEs	Y	Y	Y	Y	Y
Father Controls		Y	Y	Y	Y
Childhood Controls			Y	Y	Y
Birth Cohort FEs				Y	Y
Accident Year Cohort FEs					Y

Notes: This table reports the results from regressions of the log of wage income in 1940 on parental death. Each column gradually includes a set of covariates. The first column includes adulthood characteristics observed in the 1940 census, which includes a quadratic polynomial for age as well as adulthood county of residence and initial census year fixed effects. Column 2 adds fathers' pre-accident characteristics, containing the father's age at the time of the accident, nativity, literacy, family size, occupation-based income score, and industry of employment. Column 3 further controls for childhood characteristics consisting of the son's age at the time of the accident, birth order, the number of older siblings and sisters identified during childhood, and whether or not their childhood household included extended family. The third column also includes fixed effects for the county of residence during childhood. Columns 4 and 5 include birth cohort fixed effects and accident year cohort fixed effects, respectively. The control group mean reports the average nominal wage income among sons whose father was involved in a serious, but non-fatal mining accident. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Estimates of Losing a Father by Subgroup

	Family Support		Mining Community Size		Age at Accident
	(1)	(2)	(3)	(4)	(5)
	No Available Extended Family	No Older Siblings	Average Share	Average Size	Older at Accident
	-0.035	-0.064*	-0.021	-0.038	-0.002
	(0.028)	(0.037)	(0.037)	(0.029)	(0.025)
	Extended Family	Older Siblings	σ Above	σ Above	Young at Accident
	-0.001	-0.016	-0.028	-0.022	-0.173***
	(0.093)	(0.031)	(0.028)	(0.029)	(0.062)
N	5,526	5,526	5,526	5,526	5,526

Notes: The table presents the results of regressing the log of wage income on losing a father across several measures of subgroup heterogeneity. Each regression includes the full set of controls in the preferred specification described in Section 4.2. Each column represents a separate estimation of equation 2 with estimates of γ in the top row and $\gamma+\theta$ below. The first two columns examine differences in the main wage effects by the availability of family support. In Column 1, households are identified by whether or not they lived with extended family. Column 2 interacts the effect of losing a father with the number of older siblings in the household, such that the estimates correspond to those without siblings and those in a household with two older siblings. The third and fourth columns examine how the surrounding mining community may alter the effects of losing a father by examining the share of the community working in mining and the total number of workers in mining, respectively. Finally, the fifth column compares the effect of losing a father by the son's age at the time of their father's accident. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Estimates of Losing a Father on Other Income Measures

	(1)	(2)	(3)
Dep. Var:	Above Med. Inc.	Inc. Percentile	Non-Wage Inc.
Panel A: Average Effect			
Father Fatal Accident	-0.011	-0.438	0.034***
	(0.015)	(0.904)	(013)
Panel B: Average Effect, Young			
Fatal × Young	-0.062*	-3.276**	0.036
	(0.035)	(1.775)	(0.029)
Control Group Mean	0.571	44.461	0.124
N	5,526	5,526	5,526

Notes: This table reports the effects of parental death on three other measures of income and economic well-being. In the first column, the dependent variable is an indicator variable for whether or not the son has income above median wage income ($900) among all working-age adults born in Pennsylvania, Illinois, West Virginia, and Ohio. Column 2 shows the results for the son's percentile in the wage income distribution. Finally, the dependent variable in Column 3 is whether or not the son reported greater than $50 of non-wage income in 1940 and includes business profits, professional fees, rent collected, interest, dividends, unemployment compensation, government and private relief, pensions, annuities, royalties, and all payments-in-kind, including the consumption of one's own farm products, as non-wage, non-salary income. Panel A re-estimates β from Eq. 1 across the four alternate measures of income, while Panel B does the same for $\gamma + \theta$ from Eq. 2 and examines the effect of losing a father earlier during childhood. Control group means are presented below the point estimates and are in shares (Columns 1 and 3) and percentile points (Column 2). Standard errors clustered at the adulthood county are shown in parentheses. Each regression includes the full set of controls in the preferred specification described in Section 4.2. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Estimates of Losing a Father on Occupation Sorting

	(1)	(2)	(3)	(4)	(5)	(6)
	Mining		Skill Groups			
	Occupation	Industry	High-Skill	Semi-Skill	Low-Skill	Log(Occ. Score)
Panel A: Average Effect						
Father Fatal Accident	-0.025**	-0.029**	-0.003	-0.014	0.016	-0.010
	(0.011)	(0.012)	(0.014)	(0.017)	(0.016)	(0.011)
Panel B: Average Effect, Young						
Fatal × Young	-0.078**	-0.072***	0.008	-0.053	0.046	-0.027
	(0.031)	(0.027)	(0.026)	(0.035)	(0.035)	(0.023)
Control Group Mean	0.165	0.171	0.258	0.507	0.235	3.201
Observations	5,526	5,526	5,526	5,526	5,526	5,512

Notes: This table reports the effects of parental death on occupation sorting. In the first column, the dependent variable is an indicator for whether or not the son worked as a miner (either a mine operative or motorman), while the second column is an indicator for working in the mining industry. In Columns 3 through 5, the dependent variables are indicator variables for being employed in a high-skilled, semi-skilled, or low-skilled occupation, respectively. In the last column, the dependent variable is the log the occupational income score. Panel A re-estimates β from Eq. 1 across the four alternate measures of income, while Panel B does the same for $\gamma + \theta$ from Eq. 2 and examines the effect of losing a father earlier during childhood. Each regression includes the full set of controls in the preferred specification described in Section 4.2. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Estimates of Losing a Father on Employment

	(1)	(2)	(3)
Dep. Var:	Labor Force	Unemployed	Relief Work
Panel A: Average Effect			
Father Fatal Accident	-0.003	0.011	0.020
	(0.006)	(0.009)	(0.031)
Panel B: Average Effect, Young			
Fatal $\times$ Young	-0.007	0.063**	0.027**
	(0.012)	(0.028)	(0.012)
Control Group Mean Mean	0.930	0.097	0.131
Observations	5,526	5,526	5,526

Notes: This table reports the effects of parental death on employment. In Columns 1 through 3, the dependent variables are indicators for being in the labor force, being unemployed, and for being employed on public work relief, respectively. In Section C.3, I discuss the construction of each variable in more detail. Panel A re-estimates β from Eq. 1 across the four alternate measures of income, while Panel B does the same for $\gamma + \theta$ from Eq. 2 and examines the effect of losing a father earlier during childhood. Each regression includes the full set of controls in the preferred specification described in Section 4.2. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Parental Investment: Mothers' Outcomes

	(1) Widowed	(2) Head of Household	(3) Number of Children	(4) Labor Force
Husband Fatal Accident	0.241***	0.252***	-0.083	0.040***
	(0.031)	(0.028)	(0.116)	(0.014)
Control Group Mean	0.052	0.054	4.023	0.066
Observations	5,250	5,250	5,250	5,250

Notes: This table reports the effects of losing a husband on the outcomes of mothers. The sample consists of mothers identified by the sons of accident victims in the nearest census after they are observed during childhood. In Columns 1 and 2 the dependent variables indicate if the mother is widowed in the nearest census or reports being the head of household. In the third column, the dependent variable is the number of own children in the household. Finally, the dependent variable of Column 4 is an indicator of whether or not the mother is in the labor force. The estimates of these columns are derived from a specification that includes the fathers' pre-accident characteristics and childhood characteristics described in Section 4.2, as well as fixed effects for the initial census year observed, county of residence during childhood, and for the birth year and accident year cohorts. Standard errors clustered by childhood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Losing a Father on Childhood Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Childhood Disruptions			Educational Attainment		
Dep. Var:	Labor Force	Miner	In School	8^{th} Grade	High School	Total Education
Panel A: Average Effect						
Father Fatal Accident	0.012	-0.016	-0.006	-0.007	0.015	0.159*
	(0.010)	(0.011)	(0.010)	(0.016)	(0.013)	(0.089)
Panel B: Average Effect, Young						
Father Fatal Accident	0.006	0.005	0.000	0.036	-0.013	0.184
	(0.009)	(0.008)	(0.017)	(0.035)	(0.036)	(0.224)
Control Group Mean	0.353	0.097	0.560	0.786	0.245	12.079
N	7,124	7,124	7,124	5,489	5,489	5,489

Notes: This table reports the effects of losing a father on measures of disruptions to and accumulation of human capital. In Columns 1 through 3 estimate proxies for disruption to human capital accumulation following an accident. The dependent variables are indicators for being in the labor force, being in school during childhood, and being listed as miner in the nearest census following an accident. In these columns, the sample consists of sons of accident victims linked ten years forward to the nearest census following an accident. The estimates of these columns are derived from a specification that includes the fathers' pre-accident characteristics and childhood characteristics described in Section 4.2, as well as fixed effects for the initial census year observed, county of residence during childhood, and for the birth year and accident year cohorts. Finally, in the first three columns, standard errors are instead clustered at the childhood county level. Columns 4 through 6 estimate proxies for total human capital accumulation observed in adulthood: completing 8^{th} grade, completing high school, and total years of education. In the last four columns, standard errors are clustered at the adulthood county. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendices

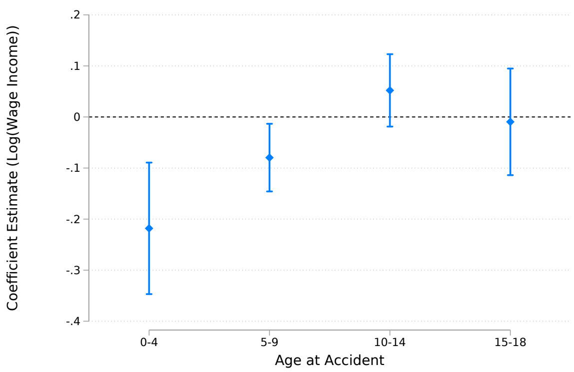
A Additional Figures and Tables

TABLE 4.—Fatal Accidents Inside and Outside of Mines

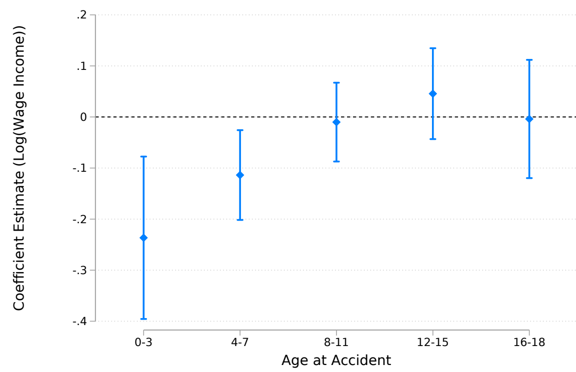
Date of accident	Name of Person	Nationality	Occupation	Age	Married or single	Number of widows	Number of orphans	Name of Mine	County	Nature and Cause of Accident in Brief
Jan. 20	Patrick McGuire,	Irish,	Driver,	24	M.	1	----	Colonial No. 4,		Fatally injured by being caught between mine cars on entry.
Feb. 6	Mike Verable,	Slavonian,	Driver,	25	M.	1	----	Paul,		Instantly killed. Supposed to have fallen off car on entry and run over.
16	Lesto Kotona,	Magyar,	Miner,	34	M.	1	2	Adelaide,		Killed by fall of roof while working on pillar stump.
17	John Malone,	American,	Elevator man,	24	M.	1	2	Colonial No. 4,		Fatally injured while trying to oil machinery in motion, contrary to orders. Outside.
March 3	George Shirrosky,	Slavonian,	Miner,	51	M.	*	5	Sumner No. 2,		Fatally injured by fall of coal at face of room while watching his fellow workman pull the coal down.
4	Andrew Kolick,	Polish,	Miner,	22	S.	----	----	Trotter,		Killed by fall of roof in pillar work while preparing to draw props. The gob slid knocking out posts.
June 17	Charles McGurke,	Irish,	Driver,	17	S.	----	----	Phillips,	Fayette,	Fatally injured by being squeezed between car and rib on narrow side of entry.
July 26	Edward Hardin,	American,	Driver,	30	S.	----	----	Lemont No. 1,		Instantly killed between car and rib on entry when car left the track.
Aug. 3	Mike Bikowski,	Polish,	Company-man	26	S.	----	----	Adelaide,		Instantly killed by fall of slate on entry. He neglected to go to a place of safety.
18	John Meson,	Slavonian,	Miner,	25	S.	----	----	Lincoln No. 1,		Instantly killed by fall of roof while drawing posts in pillar work.
31	John Franoz,	Polish,	Miner,	29	M.	1	1	Oliver No. 3,		Instantly killed by fall of slate when post was knocked out in pillar work.
Sept. 3	John Mitka,	Polish,	Miner,	37	M.	1	6	Davidson,		Fatally injured by fall of roof while drawing pillars.
28	Thomas Hutchenson,	Scotch,	Timberman,	37	M.	1	4	Colonial No. 1,		Instantly killed by fall of slate on entry while trying to secure it with cross-timbers.
21	Adolph Rottler,	German,	Miner,	51	M.	1	5	Davidson,		Fatally injured by fall of slate on haulage road.

Figure A.1: Mining Accident Record Example Page

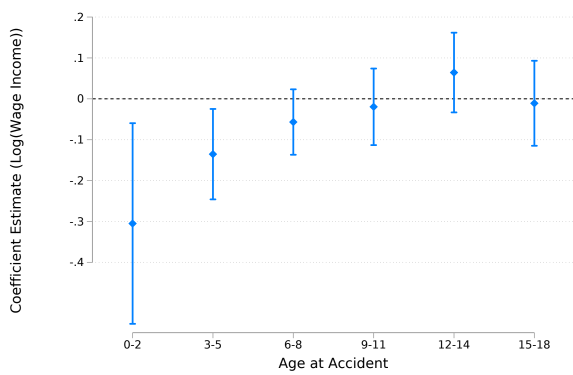
Notes: The figure presents a sample page of individual accident records published by state mining agencies. The example shows an excerpt of fatal accidents from Pennsylvania's Bituminous Coal Region during 1909 and highlights a typical record. Digitized accident records come from the Gerald E. Sherard Collection at the Colorado School of Mines and the Pennsylvania State Archives.



(a) 5-Year Groups



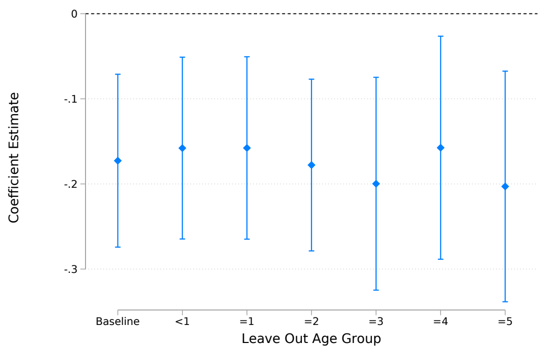
(b) 4-Year Groups



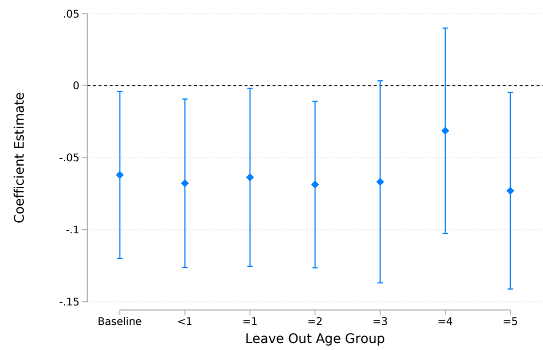
(c) 3-Year Groups

Figure A.2: Alternate Age Bin Sizes

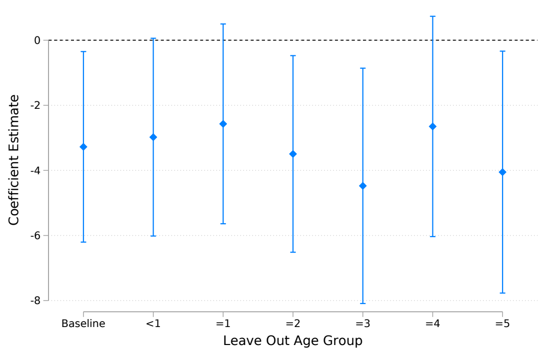
Notes: The figures plot the estimated effect of losing a father by age groups and across alternate group sizes. The figures plot the point estimates as solid points and 90 percent confidence intervals as lines. Figure A.2a presents a specification which places sons of accident victims in four groups by their age at the time of their father's accident. In Figures A.2b and A.2c, sons of accident victims are instead placed in five and six groups (each with progressively smaller bin sizes), respectively.



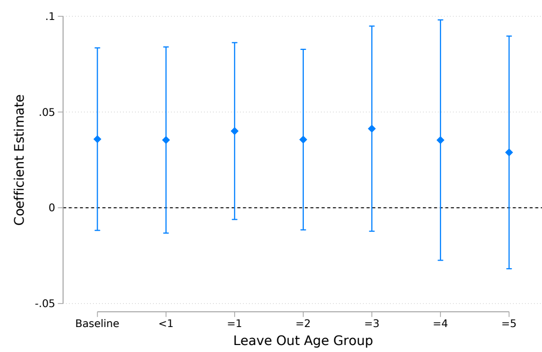
(a) Log(Wage Income)



(b) Above Median Income



(c) Income Percentile



(d) Non-Wage Income

Figure A.3: Robustness to Dropping Age Groups

Notes: This figure shows the robustness of the effect of parental loss at an early age on several measures of income. Panels A through D of the figure present the estimates (and 90 percent confidence intervals) for each measure, separately. In each panel, the leftmost estimate presents the baseline estimates reported in the main analysis. Each other estimate presents the results from omitting sons who were a specific age at the time of their father's accident.

Table A.1: Description of Accident Records

	Mean	Std. Dev.	Min	Max
Panel A: All Accident Records (N=183,536)				
Age	34.49	12.07	11	86
Birth Year	1880.76	13.09	1819	1916
Accident Year	1915.25	7.87	1900	1929
Fatal Accident	0.26	0.44	0	1
Pennsylvania	0.64	0.48	0	1
Illinois	0.23	0.42	0	1
West Virginia	0.11	0.31	0	1
Ohio	0.03	0.17	0	1
Panel B: Accident Causes (N=110,699)				
Fall of Rock	0.46	0.5	0	1
Cars and Machinery	0.35	0.48	0	1
Explosions and Gas	0.12	0.33	0	1
Other Causes	0.08	0.27	0	1

Notes: The table presents summary statistics for the state mining accident records. Panel A shows summary statistics for all records listed collected from the individual accident records of Pennsylvania, Illinois, West Virginia and Ohio. Panel B presents the three major causes of mining accidents listed in the Pennsylvania accident records, the only state for which causes were digitized. Fall of rock is any cause that mentions either fall of coal, roof, slate, or rock. Cars and Machinery contains any cause that faults runaway cars, crushed by cars, trapped by machinery, or faulty trucks or motors. Explosions and Gas contains any cause that lists dynamite explosions, faulty powder shots, delayed blasts, or dangerous gas. The remaining accidents are classified as Other and contain slips and falls (either by an individual or being hit by a tool), injuries from animals, struck by timber, suffocation due to poor ventilation, and electrocution. Digitized accident records come from the Gerald E. Sherard Collection at the Colorado School of Mines and the RG-45 Records of the Department of Mines and Mineral Industries from the Pennsylvania State Archives.

Table A.2: Summary Statistics of Fathers' Census Records

	(1) All Fathers	(2) Accident Fathers	(3) Fatal Fathers	(2)-(1) Differences	(3)-(1) Differences
Age	42.98	38.48	38.64	-4.320***	-4.121***
	[12.63]	[9.80]	[9.92]	(0.070)	(0.143)
Married	0.95	0.97	0.97	0.019***	0.012***
	[0.21]	[0.16]	[0.18]	(0.001)	(0.003)
Number of Children	2.79	3.18	3.22	0.378***	0.372***
	[1.80]	[1.96]	[1.95]	(0.014)	(0.028)
White	0.98	0.98	0.98	0.002**	-0.003
	[0.14]	[0.13]	[0.15]	(0.001)	(0.002)
Foreign Born	0.29	0.47	0.48	0.172***	0.194***
	[0.45]	[0.50]	[0.50]	(0.004)	(0.007)
Homeowner	0.45	0.36	0.34	-0.077***	-0.102***
	[0.50]	[0.48]	[0.47]	(0.003)	(0.007)
On Farm	0.21	0.1	0.1	-0.079***	-0.089***
	[0.41]	[0.29]	[0.29]	(0.002)	(0.004)
Urban Area	0.45	0.51	0.52	0.083***	0.069***
	[0.50]	[0.50]	[0.50]	(0.003)	(0.007)
Literate	0.94	0.88	0.86	-0.060***	-0.070***
	[0.23]	[0.33]	[0.34]	(0.002)	(0.005)
In Labor Force	0.96	0.98	0.98	0.025***	0.021***
	[0.20]	[0.13]	[0.14]	(0.001)	(0.002)
Occupation Score	19.82	21.13	21.34	1.059***	1.200***
	[13.37]	[10.55]	[10.51]	(0.075)	(0.152)
High Skill	0.26	0.16	0.17	-0.110***	-0.098***
	[0.44]	[0.36]	[0.37]	(0.003)	(0.005)
Semi Skill	0.19	0.5	0.49	0.284***	0.275***
	[0.39]	[0.50]	[0.50]	(0.004)	(0.007)
Low Skill	0.32	0.25	0.25	-0.077***	-0.071***
	[0.47]	[0.44]	[0.43]	(0.003)	(0.006)
Miner (Occupation)	0.06	0.42	0.41	0.339***	0.326***
	[0.25]	[0.49]	[0.49]	(0.004)	(0.007)
N	7,488,001	19,772	4,804	7,507,773	7,492,805

Notes: The table compares the mean characteristics (and standard deviation in brackets) of fathers residing in Pennsylvania, Illinois, West Virginia, and Ohio during 1900, 1910, and 1920 (Column 1) with fathers linked to mining accident records. All fathers involved in mining accidents and fathers that eventually perish in a mining accident are displayed in Columns 2 and 3, respectively. The right two columns display the differences in means between all fathers and fathers involved in mining accidents, conditional on the census year and state of residence fixed effects. The first difference column displays the differences between all fathers involved in mining accidents and all other fathers, while the second difference column displays the difference between fathers killed in mining accidents and all other fathers. Occupational skill groups follow the 1950 occupation definition of the U.S. Census Bureau. Standard errors are clustered by county of residence and are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Summary Statistics of Sons During Childhood

	Sons of Non-Fatal	Sons of Fatal	Differences
Age at Census	6.06	6.08	0.133
	[4.38]	[4.45]	(0.085)
Age At Accident	10.17	10.18	-0.053
	[4.57]	[4.58]	(0.107)
Young At Accident	0.18	0.18	0.007
	[0.38]	[0.38]	(0.008)
White	0.99	0.99	0.001
	[0.11]	[0.11]	(0.003)
Num. Siblings	3.27	3.40	0.081
	[2.11]	[2.09]	(0.063)
In School	0.50	0.51	-0.010
	[0.50]	[0.50]	(0.009)
Literate	0.22	0.23	0.018*
	[0.42]	[0.42]	(0.007)
In Labor Force	0.02	0.02	-0.000
	[0.14]	[0.14]	(0.003)
N	12,206	4,073	16,279

Notes: The table presents summary statistics for sons of fathers listed in state mining accidents that were linked to the complete count U.S. Census. The first column displays summary statistics for sons of fathers that will eventually be involved in serious, but non-fatal mining accidents. The second column displays the same statistics for sons of fathers that will eventually perish in a mining accident. Finally, the third column presents regression adjusted differences in means between the two groups of sons. Young is defined as the son being younger than primary school age at the time of their father's accident. The specification includes state and census year fixed effects. Heteroskedasticity robust standard errors are presented in parentheses and standard deviations are presented in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Estimates of Parental Death By Birth Order

	Birth Order		
	First	Second	Third or higher
Father Fatal Accident	-0.046	-0.102**	0.022
	(0.048)	(0.043)	(0.046)

Notes: The table reports the estimates of parental death on the log of wage income by birth order. The estimates stem from the regression described in Section 4.2, interacting losing a father with birth order and estimating the total effect of losing a father for each group. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Estimates of Parental Death on Occscore Measures

	(1) Occscore	(2) CW Occscore	(3) LIDO score
Panel A: Average Effect			
Father Fatal Accident	-0.01	-0.022	-0.001
	(0.011)	(0.017)	(0.010)
Panel B: Average Effect, Young			
Fatal × Young	-0.027	-0.039	-0.018
	(0.023)	(0.046)	(0.018)
N	5,512	4,323	5,188

Notes: This table reports the effects of parental death on three measures of occupation-based income score. In the first column, the dependent variable is the log of the standard occupational income score (occscore). In Column 2, I construct an alternate version of occscore in the spirit of Collins and Wanamaker (2014) that adjusts for state, industry, occupation, and urban-rural differences. Finally, in Column 3, I employ an alternate version of the typical occupational income score developed by Saavedra and Twinam (2020), which is based on occupation, industry, demographics, and geography rather than occupation alone. In Section C.3, I discuss the income score measures in more detail. Panel A re-estimates β from Eq. 1 across each dependent variable, while Panel B does the same for $\gamma + \theta$ from Eq. 2, estimating the effects of losing a father at a young age. Each regression includes the full set of controls in the preferred specification described in Section 4.2. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Other Specific Occupations

	(1)	(2)	(3)	(4)	(5)	(6)
		Laborer		Farmer		
	Non-Wage Occ.	Laborer (all)	Laborer (n.e.c.)	Farmer	Farm Labor	On Farm
Panel A: Average Effect						
Father Fatal Accident	-0.010	0.010	0.007	0.001	-0.001	-0.003
	(0.015)	(0.015)	(0.015)	(0.004)	(0.004)	(0.008)
Panel B: Average Effect, Young						
Fatal $\times$ Young	0.058	0.009	-0.001	0.022	0.010	-0.012
	(0.045)	(0.032)	(0.032)	(0.015)	(0.011)	(0.018)
Control Group Mean	0.301	0.159	0.154	0.026	0.018	0.068
N	5,526	5,526	5,526	5,526	5,526	5,526

Notes: This table reports the results from regressions of working in a specific occupation on losing a father in a mining accident. In the first column, the dependent variable is whether or not the individual worked in an occupation that earned a large share of income from non-wage sources. To calculate the variable, I take an extract of the 1950 census and calculate the share of income derived from non-wage income by each occupation. I then identify the occupations that receive more than 10 percent of income (the median share of non-wage income across all occupations) from non-wage sources. The second and third columns examine laborer occupations, either all classes of laborers or laborers not elsewhere classified. Finally, columns 4 through 6 examine proxies for farm work. Panel A re-estimates β from Eq. 1 across each dependent variable, while Panel B does the same for $\gamma + \theta$ from Eq. 2, estimating the effects of losing a father at a young age. Each regression includes the full set of controls in the preferred specification described in Section 4.2. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Robustness to Potential Confounders

	(1)	(2)	(3)	(4)	(5)
	Dry Accidents	PA Only	Accident FE	Town FE	Omit Adulthood County
Panel A: Average Effect					
Father Fatal Accident	-0.046*	-0.058**	-0.064**	-0.053**	-0.038
	(0.026)	(0.028)	(0.028)	(0.026)	(0.026)
Panel B: Average Effect, Young					
Fatal × Young	-0.184***	-0.177**	-0.181**	-0.178***	-0.170***
	(0.064)	(0.072)	(0.076)	(0.067)	(0.065)
N	5,526	3,700	3,607	5,325	5,526

Notes: This table reports the results from regressions of the log of wage income in 1940 on losing a father in a mining accident. Each column represents a modified version of the preferred specification described in Section 4.2 meant to test the sensitivity of the baseline estimates to potential confounding variation. The first column includes an indicator for whether or not the accident occurred after the enactment of a local prohibition law. The second and third columns examine only accidents that occurred in Pennsylvania, while the third column additionally includes accident-type fixed effects. Instead of childhood county of residence fixed effects, the fourth column includes childhood town of residence fixed effects from Berkes et al. (2023). Finally, the final column omits fixed effects for the son's adulthood county of residence. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Robustness to Alternative Inference Methods

	(1)	(2)	(3)	(4)
Dep. Var:	Log(Wage Inc.)	Above Med. Inc.	Inc. Percentile	Non-Wage Inc.
Fatal Accident × Young	-0.173	-0.062	-3.276	0.036
	(0.062)	(0.035)	(1.775)	(0.029)
	p-values from...			
Clustered by Adulthood County	0.005	0.079	0.066	0.215
Clustered by Childhood County	0.005	0.079	0.078	0.095
Two-Way Clustered	0.003	0.079	0.050	0.096
Clustered by Accident Cohort	0.031	0.157	0.135	0.230
Wild-Cluster Bootstrapped Ses	0.005	0.067	0.030	0.220

Notes: This table shows the robustness of the effects of parental death at an early age to alternative inference methods. Each column of the table presents the main estimate, clustered standard errors at the adulthood county level in parentheses, and p-values computed from alternative procedures for each of the measures of income. Alternative p-values are computed using (i) standard errors clustered at the adulthood county level, (ii) standard errors clustered at the childhood county level, (iii) two-way clustered standard errors at both the adulthood and childhood county levels, (iv) standard errors clustered by accident cohort, and (v) wild-cluster bootstrapped standard errors at the adulthood county level.

Table A.9: Parental Death on Mothers' Outcomes Conditional on Sons' Age

	(1) Widowed	(2) Head of Household	(3) Number of Children	(4) Labor Force
Husband Fatal Accident	0.239***	0.243***	-0.053	0.040**
	(0.032)	(0.031)	(0.126)	(0.016)
Observations	4,124	4,124	4,124	4,124

Notes: This table reports the effects of losing a husband on the outcomes of mothers, among the sample of mothers identified by sons who were younger than 18 in the nearest census following an accident. The estimates of these columns are derived from a specification that includes the fathers' pre-accident characteristics and childhood characteristics described in Section 4.2, as well as fixed effects for the initial census year observed, county of residence during childhood, and for the birth year and accident year cohorts. Standard errors clustered by childhood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B Effect of Injuries

The identification strategy relies on comparing the sons of fathers who perished in an accident to those who survived a non-fatal accident. As discussed above in Section 2, state inspectors typically reported a non-fatal accident if the victim was unable to work for a sustained period of time (typically 30 days). Some survivors, however, experienced severe injuries that kept them out of work for much longer or led to disability. If a severe, debilitating injury diminished the capacity for parental investment in the victim's household, then their son's future may have been affected as well. Put differently, the counterfactual group may be partially treated, and as a result, the effects of losing a father could be underestimated.

Testing for an Injury Effect on Sons

Estimating the effects of injuries on victims' sons requires some valid counterfactual. To test if injuries from mining accidents affected sons' long-run economic well-being, I implement a selection-on-observables design that combines exact matching and inverse propensity weights. While this approach lacks the sort of quasi-random variation afforded by the empirical strategy to estimate the effects of death, it allows me to compare sons of non-fatal accident victims to a counterfactual set of sons identified in the same state and census year with similar pre-accident household characteristics.

Specifically, I generate exactly matched groups based on the census year in which a son's father was matched, the town they lived in before the accident, and the industry their father worked in at the time of the census.[28] I then restrict the sample to include only observations within non-degenerate groups—exactly matched groups that contain at least one son whose father was listed in a non-fatal accident. To compute weights, I use a probit model to regress whether the father was listed in the non-fatal accident records on the father's characteristics and those of the household. The included characteristics are the census year and county in which the household was identified, the father's age and race, marital status, homeownership status, residence on a farm, literacy, and an exhaustive set of occupational, industrial, and birthplace indicators for the father. Estimating the model yields a predicted probability, $\hat{p}$, of being listed in a non-fatal accident. Finally, I implement inverse propensity weights by weighing counterfactual group units by $\frac{\hat{p}}{1-\hat{p}}$.

Figure B.1 plots the distribution of $\hat{p}$ for fathers injured in a mining accident and for those in the counterfactual group as solid and short-dashed lines, respectively. As one would

[28]Rather than restricting comparisons within the county where a household was identified, I use place-based clusters identified by Berkes et al. (2023) to restrict comparisons to households within the same local cluster. Doing so requires comparisons to be made within fine geographic areas.

expect, fathers in the counterfactual group, on average, are less likely to be predicted in a non-fatal accident, though the distributions have broad common support. The long-dashed line demonstrates that the differences between the two distributions disappear after weights are applied. Table B.1 similarly supports the selection-on-observables design. The first two columns display average characteristics between fathers injured in an accident and those in the counterfactual group, respectively. While injured fathers are noticeably different from those in the counterfactual, Column 3 shows the differences become small and insignificant after applying the weights and restricting comparisons within exactly matched groups.

To examine whether serious, but non-fatal mining accidents affected the adulthood outcomes of sons, I link the sons of fathers in the counterfactual group to the 1940 census. I then weighed and grouped each son by their father's corresponding propensity weight and exactly matched group. Table B.2 presents the results of regressing the adult sons' measures of economic well-being on whether their father was injured in an accident after applying both inverse propensity weights and exactly matched group fixed effects. Across each measure of income, there is little evidence that a father's injury had a lasting effect; each estimate is insignificant and not meaningfully different from zero.

Long-Term Effect of Injuries on Surviving Victims

While one may expect injuries to result in the average estimated effect of losing a father to be underestimated, the likely small share of debilitating injuries in the analysis sample suggests this was not likely the case. As highlighted in Section 2, less serious accidents were nearly 42 times as common as fatal accidents, and a contemporaneous survey of non-fatal accidents found that only ten percent of injuries were serious enough to disable a victim (Adams et al., 1932; Roberts, 1904). To directly examine how injuries affected the later employment outcomes of injured fathers, I follow a subset of fathers listed as being injured in a non-fatal mining accident to the nearest census following their injury. Doing so identifies 5,333 working-age fathers who can be found both in the nearest censuses before and after their accident.[29]

Table B.3 describes the injured fathers before and after their accident in Columns 1 and 2, respectively. Nearly all (99 percent) of the fathers were in the labor force before their accident, and two out of five worked as miners. Similarly, roughly three-quarters are listed in a semi-skilled or low-skilled occupational group. Column 3 presents before-and-after differences, net of county, census year, accident year, and birth cohort fixed effects. Two characteristics stand out following their injury. First, fathers listed in non-fatal accidents

[29] Since some fathers may have reached retirement age by the nearest census following their accident, I focus on those who were aged 20-50 in the census before their accident.

were roughly two percentage points less likely to be in the labor force. In other words, few injured fathers were unable to work due to their accident. Second, while injuries resulted in a departure from the mining industry, they did not substantially alter later employment outcomes: roughly 68 percent of injured fathers still worked in either a semi-skill or low-skill occupation group, and occupational income scores were no different following an accident. Altogether, following injured fathers and their sons suggests that while a serious, non-fatal mining accident was likely a significant short-run shock to victims and their families, the lack of long-run effects on victims themselves suggests intergenerational injury effects were small.

Table B.1: Balance Among Injured Fathers and a Matched Counterfactual Group

	(1) Non-Accident Fathers	(2) Non-Fatal Accident Fathers	(3) Weighted Differences
Age	42.98	38.44	0.102
	[12.63]	[9.76]	(0.152)
Married	0.95	0.97	0.000
	[0.21]	[0.16]	(0.002)
Number of Children	2.79	3.17	-0.004
	[1.80]	[1.96]	(0.019)
White	0.98	0.98	-0.001
	[0.14]	[0.12]	(0.001)
Foreign Born	0.29	0.47	0.005
	[0.45]	[0.50]	(0.011)
Homeowner	0.45	0.37	-0.001
	[0.50]	[0.48]	(0.006)
On Farm	0.21	0.1	-0.003
	[0.41]	[0.29]	(0.002)
Urban Area	0.45	0.5	-0.016***
	[0.50]	[0.50]	(0.005)
Literate	0.94	0.88	0.001
	[0.23]	[0.32]	(0.005)
In Labor Force	0.96	0.98	0.000
	[0.20]	[0.12]	(0.000)
Occupation Score	19.82	21.07	-0.075
	[13.37]	[10.56]	(0.088)
High Skill	0.26	0.15	-0.001
	[0.44]	[0.36]	(0.004)
Semi Skill	0.19	0.5	0.001
	[0.39]	[0.50]	(0.004)
Low Skill	0.32	0.26	0.001
	[0.47]	[0.44]	(0.003)
Miner (Occupation)	0.06	0.42	-0.001
	[0.25]	[0.49]	(0.004)
N	7,488,001	14,968	5,588,119

Notes: The table displays the mean characteristics (and standard deviation in brackets) of fathers residing in Pennsylvania, Illinois, West Virginia, and Ohio during 1900, 1910, and 1920 (Column 1) as well as fathers linked to serious, non-fatal mining accidents (Column 2). Column 3 displays the differences between the two groups by regressing each characteristic on whether the father was injured in a mining accident and after applying inverse propensity weights and conditional on exactly matched group fixed effects. Standard errors are clustered by exactly matched groups and are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Estimates of Injury Effects on Adulthood Economic Well-Being

	(1)	(2)	(3)	(4)
Dep. Var:	Log(Wage Inc.)	Above Med. Inc.	Inc. Percentile	Non-Wage Inc.
Father Injured	0.003	0.008	0.491	-0.004
	(0.017)	(0.009)	(0.508)	(0.007)
Observations	822,961	822,961	822,961	822,961

Notes: This table reports estimates of the effects of a father being injured in a non-fatal mining accident on the economic well-being of their adult sons. The sample includes the adult sons of fathers listed in a non-fatal mining accident and a counterfactual group consisting of sons whose fathers were identified in the same census year, lived in the same town, and worked in the same industry as the injured fathers. Each column presents the results of a separate regression of a measure of well-being at the top of the column on whether the father was injured in a mining accident and after applying both inverse propensity weights and conditional on exactly matched group fixed effects. Each specification also controls for individual characteristics of sons observed in the 1940 census and characteristics of their fathers and their childhood county. Individual characteristics of adult sons included whether they were in the labor force birth cohort, industry of employment, and county of residence fixed effects. Included controls for fathers are their literacy, nativity, their industry of employment, and occupational income score. Standard errors are clustered at the adulthood county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Estimates of Injury Effects on Non-Fatal Accident Victims

	Pre Accident	Post Accident	Differences
In Labor Force	0.99	0.97	-0.021***
	[0.10]	[0.17]	(0.003)
Miner (Occupation)	0.41	0.3	-0.107***
	[0.49]	[0.46]	(0.017)
Mining (Industry)	0.41	0.32	-0.089***
	[0.49]	[0.47]	(0.015)
High Skill	0.16	0.2	0.031***
	[0.37]	[0.40]	(0.012)
Semi Skill	0.49	0.4	-0.096***
	[0.50]	[0.49]	(0.015)
Low Skill	0.25	0.28	0.035***
	[0.43]	[0.45]	(0.010)
Log(Occscore)	3.15	3.14	-0.006
	[0.30]	[0.33]	(0.006)
Observations			5,333

Notes: The table displays the mean characteristics (and standard deviation in brackets) of working-age fathers listed as being injured in a mining accident who could be linked to the nearest census following their accident. The first column displays the summary statistics for these fathers in the nearest census prior to their accident, while Column 2 displays the same for the nearest census following their injury. Column 3 displays the before-and-after difference, net of county, census year, accident year, and birth cohort fixed effects. Standard errors are clustered at the county where the individual was first observed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

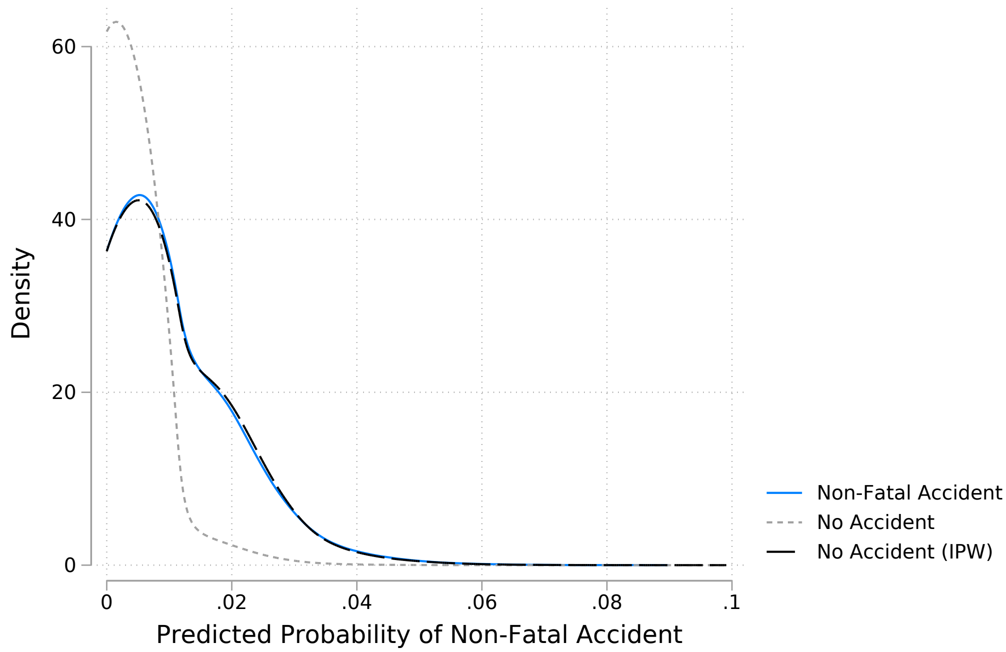


Figure B.1: Distribution of predicted propensity for non-fatal accident involvement

Notes: The figure displays the distribution of predicted propensity for having a father injured in a mining accident. The predicted propensities are estimated via a probit model, and the full set of predictors included are race, birth year, whether the father was literate, the number of children identified in the household, the census year and county the father was identified prior to the accident, and an exhaustive set of dummies for the father's occupation, industry, and place of birth.

C Data Appendix

This section provides additional details on the construction of the main datasets and variables employed in the analysis. Each section discusses the various sources for the data and the methods used to compile the raw files into their final forms.

C.1 Accident Records

Digitized mining accident records come from two primary sources: The Gerald E. Sherard collection of the Colorado School of Mines and collection RG-45 of the Pennsylvania State Archives. Each collection compiles individual mining accidents across several sources, including state mine inspection records, contemporaneous local news articles, websites dedicated to the history of mining in several local communities, and the descendants of the victims. The quality and completeness of the individual accident records vary across sources. To build the dataset, I retain only records that are reported in official state inspection reports from Pennsylvania, Illinois, West Virginia, and Ohio. Altogether, I collect 217,457 individual mining accident victims from these four states.

Individuals may appear in the records a number of times if they experienced multiple non-fatal accidents throughout their careers. Additionally, a common challenge with historical registers is that it is impossible to separately identify, since the data contain no unique administrative identifiers and individuals with common names that share a birth date. For example, Lahey Coleman born in 1900 may appear in an accident in 1922 and 1923 or even multiple times in the same year. However, historical record linking requires unique individuals, typically by name, birth year, and some level of geography (e.g., state of birth or residence).

To create a dataset compliant with the linking techniques utilized in this paper, I assemble a dataset that is unique by state, full name, and birth year. Table C.1 describes the step-by-step criteria for dealing with duplicates and creating records compliant with census linking methods. Starting with the original set of 217,457 records, I exclude any accident where the age is missing or invalid (e.g., includes non-numeric characters). I also exclude records where the accident date or full name is not complete (e.g., no last name). Since the records are compiled from several inspectors, some inspectors recorded the same victims on the same dates, I also drop duplicates in this case.

From the remaining set of records, I keep the terminal fatal accident or the latest non-fatal accident. However, this step is complicated by the fact that names (and birth years) may not be unique. In other words, the data may contain an individual who is listed in a fatal accident in 1915, and a non-fatal accident in 1916. This is particularly true for individuals with common names. As a result, if there is an individual with a fatal accident prior to a non-fatal accident at any point, they are not included in the sample.

Finally, I retain only records that correspond to accidents that occur between 1900-1929. Doing so results in a dataset of 183,536 individuals unique by name, birth year, and state in which the accident occurred. Each record corresponds to one individual and their final accident observed in the state records.

Accident Causes

The nature and cause of accidents provide a brief (and occasionally detailed) description of the incident that led to fatality or a serious, but non-fatal injury. The digitized descriptions of accidents are only available for the 110,699 records from Pennsylvania. Given the complex descriptions of some of the accidents, I group causes into five main categories: Fall of Rock, Cars and Machines, Explosions and Gas, Falls of Persons or Tools, and Other.

The first category consists of all accidents from falling rocks, coal roofs, or spoil tips. The most frequent descriptions in this category are simply "fall of rock," "fall of slate," or "fall of roof." I use a set of keywords such as *rock, slate, roof, fall of coal, collapse, rush of culm,* or *smothered by rock* to identify accidents where sudden falls of rock are the primary cause of an accident. Accidents involving collapses of roofs or other falling rocks account for nearly half of all mining accidents.

Accidents caused by Cars and Machinery contain all incidents involving mine cars, steam engines, rail cars, and various machines employed in processing coal. Such accidents are typically described simply as "by cars,""crushed by cars," "run over by cars," or involve being caught between machinery. I identify accidents caused by cars and machinery by relying on keywords such as *cars, motors, machines, breaker, engine, railcar, run over, off track,* or *faulty brake.* Over one quarter of all mining accidents are caused by cars or machinery and the majority of those accidents are the fault of cars.

Explosions, faulty blasting shots, and gas account for roughly thirteen percent of all accidents. These accidents are more likely to result in fatality and are responsible for a number of mass casualty events throughout the early 20$^{\text{th}}$ century. To identify accidents caused by explosions or gas I search for keywords like *gas, charge, explosion, blast, powder, shot* or *dynamite.* These accidents are typically described as premature or delayed blasts, flying rocks from explosions, gas ignited by fire or arcs of electricity, or collapsed coal rooms from explosions.

The fourth category contains all accidents that list falling persons or tools as the primary cause of accident. These accidents mostly include individuals slipping and falling while in the mine or fault a fall in the mine cage, a conveyance used to transport workers and supplies from the surface to the mine by way of a rope. Keywords used to identify falls include *fall of person, slipped, fell off, fell down, fell under,* or *fall of cage.* These accidents account for eight percent of all mining accidents.

Finally, accidents not elsewhere classified are included in the Other category. A catch-all, this category consists mostly of accidents cause by animals, suffocation, drowning, or electrocution and accounts for the remaining set of accidents.

Identifying accident causes using keywords is an imperfect measure of assigning the cause of accident since individual accident causes may be identified as having multiple sources. In some cases, I can reasonably assign fault to cases with multiple sources. For example, "fall of rock from explosion" can safely be assigned to the category of explosions and not the sudden fall of rock. However, some 2,000 incidents have multiple causes that cannot be isolated as the primary cause. For example, "by car and rock" has no clear primary cause. In these few cases, I allow the incident to have multiple causes.

C.2 Linking Process

I employ commonly used techniques to historical record linking. Specifically, I employ a refined version of the Ferrie (1996) algorithm developed by Abramitzky et al. (2021). To begin, I resolve discrepancies in naming conventions between census data and the accident records by systematically standardizing first names and replacing commonly abbreviated names with full names. For example, nicknames and shorthand are common in the accident records (E.g. JNO instead of John or WLM instead of William).[30] I define candidate links based on age, state of residence, and similarity scores computed by first and last name. Similarity scores are Jaro-Winkler string distances which give a measure of the similarity between two strings, placing more weight on similar characters at the beginning of the string. A similarity score of 0 represents a candidate whose name exactly matches, while a score of 1 represents two names with no shared characters.

The algorithm relies on fixed information, typically place of birth, to refine the pool of potential links between sources. Since the accident records do not report place of birth, I instead (i) assume individuals listed in a given state's accident records lived in that state at the time of the census and (ii) search for individuals in the nearest prior census. For example, if an individual accident occurred in Pennsylvania during 1904, I search for that accident victim among Pennsylvania residents in the 1900 U.S. Census.

Similar to the procedure outlined in Abramitzky et al. (2021), I implement the following algorithm:

[30] I do not rely on phonetic name cleaning algorithms like Soundex or NYSIIS since Bailey et al. (2020) show that use of phonetic algorithms may increase false positive links.

1. For each observation in dataset A, I define a set of candidate matches in dataset B. An observation in dataset B is a candidate match for an observation in A if:
    - It has the same state of residence.
    - The birth year is within plus/minus two years of the reported birth year in dataset A.
    - The initials of first and last names are the same between candidates.
2. For each pair of candidate matches, I compute the Jaro-Winkler score for the first and last name. I consider a candidate pair matched on names if the total Jaro-Winkler score is less than or equal to 0.10. After computing name similarity scores there are three possibilities:
    - There is no matched candidate for a given observation in dataset A. In other words, there are no records in dataset B that have sufficiently similar (Jaro-Winkler score ≤ 0.10) names in the same state and within the age window.
    - The observation in dataset A has a unique matched candidate in dataset B. In this case, the observations are considered linked and are included in the next phase of the analysis.
    - The observation in dataset A has more than one record that is considered matched on names and within plus/minus two years of birth. In this case, I take the candidate with the smallest difference in reported age as the linked record and include it in the next phase of the analysis. If there are still more than one records with the smallest difference in reported age, there is no unique match and the records are not used.
3. Next, I repeat the above steps, matching observations from dataset B to A (instead of from A to B). The sample that continues on to the next phase of the analysis is the intersection between observations that match uniquely from dataset A to dataset B and those that match uniquely from dataset B to dataset A.

Linking the accident records to the complete count census years provides a total of 53,510 matches from the 183,536 accident records. The match rate of 29.1 percent is within the standard matching rate in the historical census linking literature which frequently finds match rates between 20-30 percent. Of the 53,510 accident victims matched to the census, I identify 19,772 victims who were listed as fathers at the time of enumeration. From these heads of household, I collect 16,279 native-born sons to link forward to their adulthood outcomes in 1940. I focus only on native-born sons for two reasons. First, I focus on sons since linking daughters presents the additional challenge of marriage. Since women customarily change their surname after marriage, it is not possible to link daughters unless one possesses information about renaming post-marriage. Second, I focus only on native-born children since nearly 98 percent of those identified as children of accident victims are born in their father's state of residence.

In the next step, I link sons of mining accident victims forward to the 1940 Census. For the 16,279 native born sons, I repeat the record-linking algorithm described above; I link each group of sons found in their respective census year forward to the 1940 Census. I then focus only on sons who were younger than 18 at the time of their father's accident since sons who reach young adulthood are less likely to be co-residents during the accident. Moreover, children who were older than 18 and still co-resident with their parents may differ in unobserved ways, and these differences may reflect some sort of underlying qualities that matter for long-run outcomes. Focusing on adult sons who were younger than 18 at the time of their father's accident yields 5,624 men observed in 1940 with valid responses in all outcomes and control variables. Of these adult children of accident victims, 23.8 percent lost their father due to a mining accident.

Selection into matching presents a potential bias if records that match fail to capture the characteristics of the average miner involved in an accident. To probe this potential bias, I compare records that match to the 1900, 1910, and 1920 complete count census with records that fail to find a match in any census year. Table C.2 presents the mean characteristics of matched and not-matched accident records as well as differences between the two means. Accident victims that successfully matched to the complete count census were younger and less likely to perish as a result of their mining accident. Figures C.1a and C.1b plot the distribution of age at the time of accident and birth year of linked and not linked accident victims. These figures provide further evidence that matched and not-matched victims are not substantially different. Finally, Figure C.2 displays the differences between matched and not-matched accident victims standardized by the mean and standard deviation of the not-matched victims in gray. The figure highlights the magnitudes of the differences; differences are clustered around less than ten percent of a standard deviation and are precisely estimated.

Still, to account for potential biases due to record linking, I compute inverse propensity weights (IPWs) for each individual listed in the accident records. To do so, I first estimate a probit model where the dependent variable indicates whether the individual was matched to a census record. In this model, I predict the probability of linking with the string length of the victim's first and last names, whether or not they perished in an accident, and accident year, accident month, state of residence, and birth year fixed effects. I then predict the conditional probability of being linked and construct IPWs as suggested by Bailey et al. (2020). Weighted differences are shown in green in Figure C.2 and are smaller in magnitude than unweighted differences. Finally, Table C.3 compares the main estimates of the paper to estimates weighted by the computed IPWs. The weighted and unweighted estimates are similar in both magnitude and precision, suggesting that selection into the linked sample poses no significant threat to the main identification strategy.

C.3 Complete Count Census Files

I rely on full-count data files for the 1900, 1910, 1920, and 1940 U.S. Census. The Census files are provided by IPUMS-USA via a restricted-use licensing agreement. Importantly, the restricted access files contain full names of individuals in the U.S. Census. In this section, I describe the variables employed in the main analyses and their construction.

Income Measures

In order to construct the log of annual wage income reported in 1940, I rely on an IPUMS Census variable, *incwage*, which reports each respondent's total pre-tax wage and salary income for the previous year.[31] Sources of income include wages, salaries, commissions, cash bonuses, tips, money received from public emergency work, and other money income received from an employer. Importantly, this variable measures only money received as an employee during the previous year and does not capture income from assets or other factors of total compensation. In the main analysis, I exclude any respondent who did not report positive annual wage income and re-code top-coded values as $5,000.

While the 1940 Census was the first to report wage income, enumerators were not instructed to measure income from other sources. Instead, respondents were simply asked to report if they received at least $50 of non-wage income. In the 1940 census, non-wage income was meant to capture business profits, professional fees, rent, interest, dividends, pensions, annuities, royalties, and all payments-in-kind. Non-wage income also included unemployment compensation, as well as government and private relief. I use this response to create a dummy variable for whether or not the respondent reported earning at least $50 of non-wage income.

Occupation-based Income Score

Before the 1940 census, data on individual wage income was unavailable in the U.S. Census. As a result, researchers seeking to understand labor market outcomes of individuals relied

[31]For further details, see https://usa.ipums.org/usa-action/variables/INCWAGE#codes_section

on occupation-based indices of economic outcomes. Arguably the most popular measure is the occupational-based income score (occ. score), which measures the median income of an occupation in 1950.

While the classic occ. score provides a reasonable proxy for occupational standing rather than individual earnings, the measure may miss important heterogeneity due to geographic and demographic characteristics.[32] Instead of the typical occ. score, I also employ alternative measures suggested by Collins and Wanamaker (2014) and Saavedra and Twinam (2020).

To construct occupation scores in the spirit of Collins and Wanamaker (2014). I retrieve a 1% sample of the 1950 census and examine the average total income of working-age men (aged 20-60) within cells of state, urban/rural status, industry and occupation, and employment status (I do not stratify by race, since 99 percent of the accident sample is White). I then take the log of this adjusted income score and match it to the analysis sample by the specified cells. I also employ a LASSO-adjusted measure of occupation score generated by Saavedra and Twinam (2020). They implement a machine learning approach to construct a new income score based on industry, occupation, geography, and demographics.

Occupation Codes

In 1940, Census enumerators were asked to record respondents' occupations and industry of employment. The responses generated tens of thousands of occupations and industries with limited guidance for comparability. IPUMS has aggregated most of these ambiguous occupations into roughly 200 granular and nine broad categories.[33] The broad categories, along with some example occupations from each class, are:

1. Professionals and Technical Workers: professors, engineers, and healthcare providers.
2. Farmers: farm managers, owners, and tenants.
3. Managers, Officials, and Proprietors: store managers, public administration inspectors, building managers.
4. Clerical and Kindred: clerks, bookkeepers, office machine operators.
5. Sales Workers: insurance agents and brokers, and salesmen.
6. Craftsmen: blacksmiths, carpenters, and tailors.

[32] For instance, Bailey and Collins (2006); Collins and Wanamaker (2014); Inwood et al. (2019) and Saavedra and Twinam (2020) all highlight that occupational income measures that incorporate geographic and demographic characteristics may better reflect the income standing of individuals within occupations.

[33] For details, see https://usa.ipums.org/usa-action/variables/OCC1950

7. Operatives: mine operatives and laborers, sawyers, and truck drivers.

8. Service Workers: housekeepers, bartenders, and wait staff.

9. Laborers: farm laborers, lumbermen, day laborers.

From these nine broad occupation categories, I further group occupations into high- semi- and low-skilled occupations. High-skilled occupations are comprised of professionals, managers, and craftsmen. Semi-skilled occupations are comprised of sales workers, clerical workers and operatives. Low-skilled occupations include farmers, service workers, and laborers.

Employment

In Section 5, I rely on a series of variables in order to examine the effects of early parental loss on employment outcomes both along the extensive and intensive margins. These variables measure whether respondents were part of the labor force, unemployed, worked a public relief job, and the number of weeks employed during the previous year. This section briefly describes the construction of these variables and the source variables in the 1940 Census records made available through IPUMS (Ruggles et al., 2021).

The primary variable used to construct measures of employment is the respondent's employment status. This categorical variable indicates whether the respondent was part of the labor force (i.e. working or unemployed and seeking work) and, if so, whether the respondent was currently at work. From this categorical variable, I construct simple indicators that mark if a respondent was part of the labor force and if a respondent was currently unemployed.

Importantly, Census enumerators in 1940 were instructed to ascertain whether respondents were either at work on, or assigned to, public emergency work projects conducted by the Work Progress Administration, National Youth Administration, Civilian Conservation Corps, or through some state or local work relief agency. Such respondents were classified as being "on public emergency work." For these respondents, I create an indicator variable that identifies workers who reported their employment status as being affiliated with a public emergency work project.

It is worth noting the degree of misclassification and under-counting surrounding emergency work in the 1940 Census. For instance, the numbers of public emergency workers reported in the United States was just over 2.5 million, while the number recorded by the combined payrolls of the Federal agencies dedicated to public work was over 3.3 million (Census Bureau, 1946). The Census Bureau noted confusion on the part of both enumerators and respondents regarding the classification of certain types of emergency work, and the reluctance of some to report their status as on public work relief. Indeed, the most common type of under-counting occurred when those workers on emergency work listed themselves as "at work" rather than working on a public emergency project.

Finally, enumerators asked respondents to count all weeks over the last year in which any work was done. Enumerators were instructed to calculate the number of full-time equivalent weeks the respondent worked for profit, pay, or as an unpaid family worker during the previous year. For workers, total weeks at work included paid vacations and other paid absences.

C.4 Placebo Estimates

Statistical inference is further complicated in this setting since models with relatively few treated units and clusters can lead to improper inference (Cameron et al., 2008; Ferman and Pinto, 2019; MacKinnon and Webb, 2017). It is common to assume that the error term is correlated within clusters, but uncorrelated between them, yet test statistics based on a cluster-robust variance estimator tend to over-reject the null hypothesis when the number of clusters is small. While the wild cluster bootstrap estimator proposed by Cameron et al. (2008) can often lead to more reliable inference, MacKinnon and Webb (2017) show that it can still lead to improper inference if the number of treated clusters is relatively small. In this particular setting, one may share this concern. For instance, while constructing standard errors at the adulthood county of residence level yields 594 clusters, the relatively few sons who experienced the death of a father during early childhood are represented in just 114 of the clusters.

To address this concern, I follow Chetty et al. (2009) and conduct a non-parametric placebo test of the effect of early parental loss on adulthood wage income. To do so, I link nearly 1.6 million sons born in Pennsylvania, West Virginia, and Ohio whose fathers were not involved in a mining accident. Of these non-accident sons, I randomly draw a sample of equal size to the actual analysis sample above and assign placebo indicators to mimic treatment group assignment. Specifically, I randomly assign a placebo indicator for parental loss and being young at the time of an accident to 25 percent and 18 percent of these sons, respectively.[34] Following randomization, I estimate variants of Eq. 2 where the $FatherFatal_i$ and Z_i indicator variables are replaced with the alternative, placebo indicator variables. I repeat this exercise 1,000 times, which yields a distribution of $\widehat{\gamma+\theta}$. The distribution of placebo estimates represents the sampling distribution of $\widehat{\gamma+\theta}$. To compute the p-value

[34] In Table A.1, I show that roughly one-quarter of victims listed in the mine accident records perished as a result of their accident. Additionally, Table A.3 shows that 18 percent of children were less than primary school age at the time of their father's accident.

associated with the null hypothesis that the estimated income difference among those who lost their parent during early childhood is no different from the placebo estimates, I calculate the percentile of the actual estimate in the distribution of placebo estimates. Since this exercise makes no parametric assumptions about the variance-covariance matrix, nor does it suffer from biases arising from small numbers of treated clusters, I view this as an alternative and conservative approach to statistical inference.

The first panel of Figure C.3 shows the empirical cumulative distribution functions of placebo estimates (as gray dots), as well as the actual estimates (as a blue diamond). The figure highlights that the actual estimate of parental loss on adulthood income is much larger in absolute value than the placebo estimates and remains statistically significant at conventional levels. The remaining panels of the figure similarly show the empirical cumulative distribution function of placebo estimates for the other measures of income.

Table C.1: Criteria for Analysis of Accident Records

Selection Criterion	Retained Records
Original set	217,457
Drop if age listed in accident records is missing	203,223
Drop if missing full accident date or full name	200,080
Drop duplicated records within the same date	192,556
If fatal comes before non fatal, drop all. If fatal is the final record, keep the fatal accident. Otherwise, take the latest accident.	191,231
Keep accidents between 1900-1929	183,536

Notes: The table describes the criteria used to assemble the unique accident records from the full set of state mine inspectors' reports. The original set of data includes fatal and non-fatal mining accidents from Pennsylvania, Illinois, West Virginia, and Ohio. To create a set of records compliant with historical record linking, the data is transformed into a set of individuals unique by state of accident, full name, and year of birth.

Table C.2: Comparing Matched and Not Matched Records

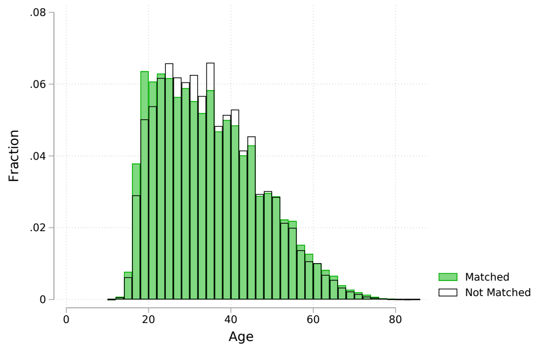
	All	Matched	Not Matched	Difference
Age	34.46	34.36	34.5	-0.134**
	[12.10]	[12.51]	[11.93]	(0.063)
Birth Year	1880.79	1881.69	1880.42	1.277***
	[13.12]	[13.68]	[12.86]	(0.069)
Accident Year	1915.24	1916.05	1914.91	1.143***
	[7.87]	[7.91]	[7.83]	(0.041)
Fatal Accident	0.26	0.24	0.26	-0.024***
	[0.44]	[0.43]	[0.44]	(0.002)
Pennsylvania	0.64	0.66	0.63	0.031***
	[0.48]	[0.47]	[0.48]	(0.002)
Illinois	0.23	0.23	0.23	-0.001
	[0.42]	[0.42]	[0.42]	(0.002)
West Virginia	0.11	0.08	0.12	-0.033***
	[0.31]	[0.27]	[0.32]	(0.001)
Ohio	0.03	0.03	0.03	0.003***
	[0.17]	[0.18]	[0.17]	(0.001)
Fall of Rock	0.45	0.43	0.46	-0.028***
	[0.50]	[0.50]	[0.50]	(0.003)
Cars or Machinery	0.29	0.31	0.28	0.033***
	[0.45]	[0.46]	[0.45]	(0.003)
Explosion or Gas	0.13	0.11	0.13	-0.019***
	[0.33]	[0.32]	[0.34]	(0.002)
Fall of Person or Tools	0.08	0.09	0.08	0.007***
	[0.27]	[0.28]	[0.27]	(0.002)
Other Causes	0.05	0.05	0.05	0.006***
	[0.22]	[0.23]	[0.21]	(0.001)
N	183,536	53,510	130,026	183,536

Notes: The table shows summary statistics for accident records that successfully linked to a record in the complete count census in 1900, 1910, and 1920 and for records that fail to link. Columns 1, 2, and 3 each present the mean and standard deviation in square brackets) of several variables for all accident records, accident records linked to the census, and for those records that fail to link, respectively. The right most column displays a difference in means between not matched and matched records along with heteroskedasticity robust standard errors in parentheses. Digitized accident records come from the Gerald E. Sherard Collection at the Colorado School of Mines and are collected from 1900 to 1929. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

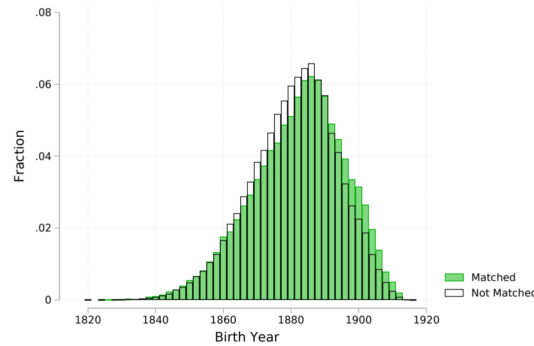
Table C.3: Weighted Estimates of Parental Death on Income Measures

	(1)	(2)	(3)	(4)
Dep. Var:	Log(Wage Inc.)	Above Med. Inc.	Inc. Percentile	Non-Wage Inc.
Panel A: Average Effect				
Father Fatal Accident	-0.052*	-0.018	-0.839	0.037***
	(0.028)	(0.015)	(0.974)	(0.013)
Panel B: Average Effect, Young				
Fatal×Young	-0.197***	-0.077**	-3.979**	0.043
	(0.068)	(0.036)	(1.918)	(0.031)

Notes: This table reports the results from regressions of parental death on measures of income in 1940, weighted by propensity weights generated to account for selection into the linked sample discussed in Section C.2. Panel A re-estimates β from Eq. 1 across each dependent variable, while Panel B does the same for $\gamma + \theta$ from Eq. 2, estimating the effects of losing a father at a young age. Standard errors clustered at the adulthood county are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$



(a) Histogram of Age by Matched



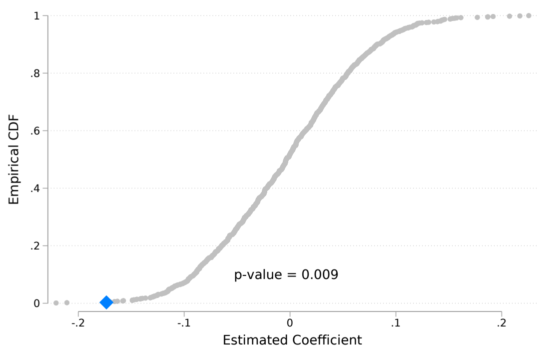
(b) Histogram of Birth Year by Matched

Figure C.1: Histograms (Matched Vs. Not Matched)

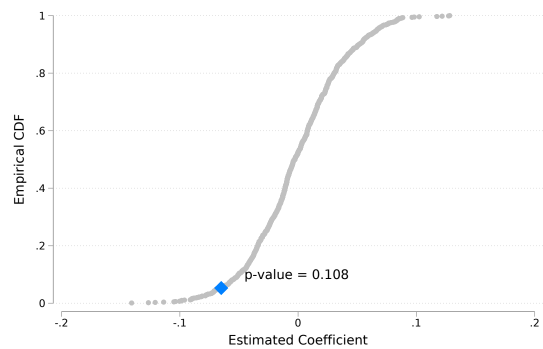
Notes: Figure presents the distributions of age at the time of accident and birth year within the individual accident records by whether or not the record matched to the 1900, 1910, or 1920 Census. The distributions of matched accident victims are presented by green bars while the distributions of non-matched accident victims are presented by hollow bars. Individual accident records come from the Gerald E. Sherard Collection at the Colorado School of Mines.

Figure C.2: Difference in Means, SD of Not Matched

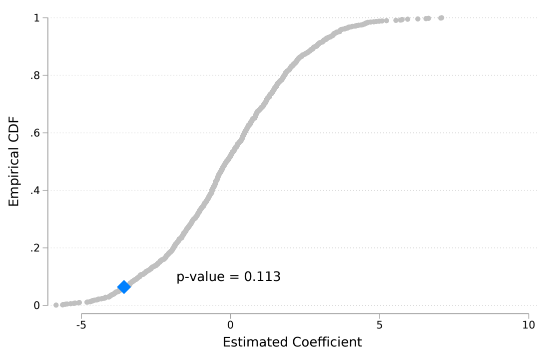
Notes:The figure shows comparisons between matched and not matched records. I standardize variables using the mean and standard deviation of the not matched group. I plot the standardized coefficient (along with 95 percent confidence intervals) from a regression of each characteristic on a dummy for accident victims that match to a census record in 1900, 1910, or 1920 in gray. In green, I plot standardized differences weighted by inverse propensity weights suggested by Bailey et al. (2020). Individual accident records come from the Gerald E. Sherard Collection at the Colorado School of Mines.



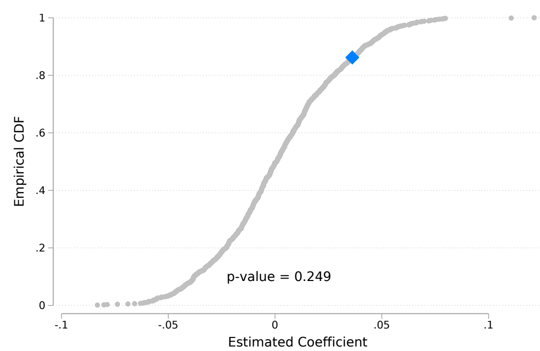
(a) Log(Wage Income)



(b) Above Median Income



(c) Income Percentile



(d) Non-Wage Income

Figure C.3: Distributions of Placebo Estimates

Notes: The figure presents the distributions of placebo estimates. Panels A through D show the empirical cumulative distribution functions of placebo estimates (as gray dots), as well as the actual estimated $\gamma + \theta$ from Eq. 2 (as a blue diamond), separately for each measure of income. Placebo estimates are obtained by estimating 1,000 variants of Eq. 2 and are discussed in more detail in Section C.4.