Train to Opportunity: the Effect of Infrastructure on Intergenerational Mobility*

Julián Costas-Fernández[†] José-Alberto Guerra[‡] Myra Mohnen[§]

February 14, 2022

Latest version

Abstract

Can transport infrastructure promote long-term labor opportunities and break the occupation tie between parents and their children? This paper estimates the causal effect of access to the railroad network on intergenerational mobility in nineteenth century England and Wales. By linking individuals across the full-population 1851, 1881 and 1911 censuses, and geolocating addresses, we determine how proximity to the nearest train station affected the occupation mobility between fathers and sons. To address the non-random access to the railroad, we create a dynamic hypothetical railroad based on geographic features. We find that sons who grew up approximately 5 km closer to a train station were 11 percentage points more likely to work in a different occupation than their father and 5 percentage points more likely to be upward mobile. The majority of the effects are driven by changes in local labor opportunities.

Keywords: intergenerational mobility, infrastructure, spatial mobility
JEL codes: H54, J62, N13

*We thank the Economic History Society for financial support. We also thank the UK Data Service and the Cambridge Group for the History of Population and Social Structure for kindly sharing their data. We are grateful to Jonathan Créchet, Jason Garred, Rick Hornbeck, Louis-Philippe Morin, Stephen Redding, Andreas Uthemann, Felipe Valencia Caicedo and seminar participants at the RATA University of Los Andes, University of Ottawa, VATT Applied Economics Seminar Helsinki, EHS, RES, Vrij University Amsterdam, University of Bath, CReAM-UCL, University of Middlesex, CEA Summer Meeting, IFN Stockholm, NASMES, NBER Urban Economics, CEDE University of los Andes, EEA, LANE HOPE, Queen's University, UQAM, NBER Economic Mobility, AMSE-BDF-CEPR Social Mobility, and University of Cologne.

[†]CReAM, University College London; j.costas-fernandez@ucl.ac.uk
[‡]Assistant Professor, Department of Economics, Universidad de los Andes; ja.guerra@uniandes.edu.co
[§]Assistant Professor, Department of Economics, University of Ottawa; Research Associate, Centre for Economic Performance, LSE; mmohnen@uottawa.ca

1 Introduction

Steam locomotives were invented in Britain in the early nineteenth century and used for railroad transport over the next century. By 1914, Britain had built a railroad network with about 32,000 km of tracks. By providing the transport of freight and passengers more quickly and cheaply than ever before, the railroads brought new economic opportunities. Rostow (1959) famously stated that "the introduction of the railroad has been historically the most powerful single initiator of take-offs".

While the economic impacts of transport infrastructure have received significant attention, the focus has generally been on aggregate or cross-sectional individual-level outcomes.[1] We know much less about how access to transport infrastructure affects individuals' economic opportunities, especially in the long-run. Transport infrastructures can improve the economic opportunity of individuals by connecting residents to job opportunities further away and/or creating better options locally. In the long-run, this has the potential to break the link between parents' economic status and their children's outcomes, that is, to increase intergenerational mobility.

In this paper, we estimate the causal effect of individuals' access to the railroad network on intergenerational mobility. We exploit the expansion of the railroad network in nineteenth century England and Wales. We create a new and unique dataset of close to 1 million father-son pairs for which we observe intergenerational occupation mobility and their proximity to the railroad network. Thanks to the newly digitized full-population 1851, 1881 and 1911censuses of England and Wales (Schürer and Higgs, 2020), we identify individuals across censuses using the linking method proposed by Abramitzky, Mill and Pérez (2019). This allows us to measure intergenerational mobility by comparing the occupation of sons as adults and their fathers' during their youth. Occupation ranking allows us to determine upward and downward mobility. We geographically locate individuals down to the street level based on the address of residence. This dataset permits the analysis of intergenerational mobility of a large and representative sample at a more geographically disaggregated level than was previously feasible. By overlaying the digitized railroad network (Alvarez, Bogart, Satchell, Shaw-Taylor and You, 2017), we measure individual access to the railroad network as the geographical proximity between the place of residence and the nearest train station during

[1] The evaluation of transport infrastructure that has largely focused on aggregate outcomes such as regional trade (e.g., Donaldson, 2018; Faber, 2014), agricultural trade and income (e.g., Donaldson and Hornbeck, 2016), urbanization (e.g., Baum-Snow, 2007; Duranton and Turner, 2012), economic growth (e.g., Banerjee, Duflo and Qian, 2020), literacy rate (e.g., Michaels, 2008), and high school enrollment.

their youth.

The access to the railroad network is likely correlated with demand for trade, migration, local resources, and/or cost of land. Railroad companies wanted to connect commercial centres at the lowest cost, which raises the concern that connected locations were on a different growth trajectory. Individuals' characteristics such as wealth or preferences likely determined their choice of residence and their intergenerational mobility patterns. To address the endogeneity in the proximity to the railroad network, we create a time series of hypothetical railroad network based solely on geographic features, ignoring demand-side concerns for railroad companies and location decisions for households. This allows us to isolate the portion of the variation in the proximity to the railroad network that is attributable to exogenous cost considerations. We use the proximity to the nearest line in the hypothetical network as an instrument (e.g., Alvarez et al., 2017; Banerjee et al., 2020; Chandra and Thompson, 2000; Faber, 2014; Lipscomb, Mobarak and Barham, 2013; Michaels, 2008). The identification strategy exploits the fact that individuals located along a counterfactual convenient route were more likely to be better connected. In addition, we control for potential correlation between location and economic characteristics due to history and/or sorting. We compare the intergenerational occupation mobility of individuals who grew up closer to a railroad station to those who grew up further away, conditional on county and census year fixed effects, and a set of control variables including proximity to historical centres, historical travel routes, and household characteristics.

We find that growing up closer to a train station led to a significant break between the occupations of fathers and sons and increased upward mobility. Sons who grew up one standard deviation (roughly 5km) closer to a train station were 11 percentage points more likely to work in a different occupation than their father. They were also 5 percentage points more likely to be upward mobile (i.e. work in an occupation ranked one standard deviation higher than their fathers). The results are driven by significant transition out of farming activities and into industrial and commercial activities. This had distributional consequences, particularly benefitting sons from the bottom of the occupational ranking. For sons from the bottom of the occupational ranking, better access to the railroad network were significantly more likely to be upward mobile and less likely to be downward mobile. For sons from the middle of the occupational ranking, better access to the railroad network were more likely to be in a different occupation than their father. These results are robust to a wide range of controls, specifications, and robustness checks.

Did the connection to the railroad network promote intergenerational mobility by facilitating spatial mobility? Or did it improve local labor market opportunities? We decompose the effect of growing up closer to the railroad network on intergenerational mobility into three channels: the change in the ease of spatial mobility, the change in the relative benefit from moving, and the change in local labor market opportunities. Our decomposition exercise reveals that local opportunities account for majority of the effect of the access to the railroad network on intergenerational mobility accounting for 78% of occupational mobility and 97% of upward mobility. In contrast, spatial mobility and the change in the relative benefit from moving only account for a small fraction. When examining spatial mobility, we find that better connected sons were 15 percentage points more likely to move away from the county where they grew up. To estimate the return to spatial mobility, we compare sons who moved away from their childhood county to their brothers who stayed put. This enables us to account for the selection into mobility across households (Abramitzky, Boustan and Eriksson, 2012). We find that the railroad decreased the relative benefits from moving. This comes from the fact that the train brought new labor opportunities to residents by changing the local economic landscape and/or expanding the labor market thanks to the possibility of commuting (Heblich, Redding and Sturm, 2020). Better connected parishes experienced urbanization, industrialization and change in their social structure. We find that the railroad network allowed people to flock to cities, industries to expand, and a new class of wealthy entrepreneurs to form. Consequently, the railroad altered the social structure with higher local occupational ranks and inequality.

There is significant evidence across countries that lower-income populations tend to suffer from restricted transport options (e.g., Chetty and Hendren, 2018; Chetty, Hendren, Kline and Saez, 2014). The poor access to transport options limits access to jobs, educational institutions and health facilities, which in turn can lead to "poverty traps". There is a long-standing debate regarding the approaches to combat inequality and uneven development. "People-based" policies aim to increase the opportunities by targeting directly low-income households (e.g. Moving to Opportunity or Earned Income Tax Credit) while "placed-based" strategies aim to increase opportunities by targeting underperforming neighborhoods (e.g. Empowerment Zone program or European Union Structural Funds). Large transport infrastructure projects have recently been proposed to specifically tackle the rise in inequality in opportunities.[2] Our results suggest that, at least in nineteenth century England and Wales,

[2]For instance, President Biden's $2 trillion "Build Back Better" proposal states that it will spark "the second great railroad revolution" by connecting workers to jobs, and spurring investment in communities that will be better linked to major metropolitan areas `https://joebiden.com/clean-energy/`. The high speed railway linking up London, the Midlands, the North and Scotland (HS2) is expected to cost between

transport projects created local economic opportunities and improved intergenerational mobility.

This paper contributes to several strands of the literature. There is a vast analytical and empirical literature has been concerned with the effects of infrastructure development on income growth, productivity and welfare (see Redding and Turner (2015) for a summary). Our results confirm previous findings that the construction of railroads led to increase income (Donaldson, 2018), migration (Morten and Oliveira, 2014; Sequeira, Nunn and Qian, 2020), literacy (Chaudhary and Fenske, 2020), regional disparities (Chatterjee and Turnovsky, 2012), the number of factories (Atack, Haines and Margo, 2020), and accelerated urbanization and city growth (Baum-Snow, 2007; Duranton and Turner, 2012). There is a general consensus amongst economic historians that railroads brought significant benefits to British society by fostering economic growth, stimulating population increase and facilitating urbanization (Alvarez et al., 2017; Baker, 1971; Bogart, Xuesheng, Alvarez, Satchell and Shaw-Taylor, 2020). In contrast to the previous literature that has used data on connected places or its residents, we exploit individual longitudinal data thereby allowing us to track geographic mobility. This allows us to find highly localized and heterogeneous effects. Living even 5km closer to the train station has a significant effect on the economic opportunities of an individual. The railroad did not benefit all residents equally. It particularly benefitted sons from lower occupational ranking. Moreover, our decomposition exercise show that the majority of the changes in intergenerational mobility patterns are driven by changes in local labor market opportunities.

We also contribute to the literature documenting intergenerational mobility. Researchers have used marriage registrations (Miles, 1999), family histories (Prandy and Bottero, 2000), surnames (Barone and Mocetti, 2021; Björklund and Jäntti, 1997; Clark and Cummins, 2015; Güell, Rodríguez Mora and Telmer, 2015), and first names (Olivetti and Paserman, 2015). Using a subsample of census, Long (2013) show that, during the nineteenth century in Britain, social mobility is greater than what was previously documented once life-cycle patterns are accounted for. Thanks to newly digitized full-population censuses, we link close to 1 million individuals across censuses with match rate of 42-49%. This allows us to document intergenerational mobility on a larger set than was previously possible. By locating individuals down to the street level, we are the first to uncover striking patterns of spatial clustering of intergenerational mobility at very disaggregate level.

£65 and £88 billion and lists as one of its aim to bring jobs and investment to the Midlands and North `https://www.hs2.org.uk/why/connectivity/`.

While the literature documents differences in intergenerational mobility across regions within countries and over time, the factors that determine changes and differences in intergenerational mobility are not yet well understood. Many public interventions affect intergenerational mobility such as tax schemes (Chetty and Hendren, 2013; Piketty, 2000), education (Machin, 2007; Milner, 2020), welfare receipt Levine, Zimmerman et al. (1996), and neighborhood influences (Chetty and Hendren, 2018; Guerra and Mohnen, 2020; Long and Ferrie, 2013). These factors shape access to physical capital and accumulation of human capital. Alesina, Hohmann, Michalopoulos and Papaioannou (2021) find that colonial investments in the transport network and missionary activity are associated with upward mobility. Perez (2017) uses the expansion of railroad network in the nineteenth century Argentina to look at how the reduction in transport costs affected the economic outcomes of parents and children. He finds that once a district got connected to the railroad, adults remained in farming activities whereas children moved out of farming towards white-collar and skilled blue-collar jobs. We distinguish ourselves from these papers in terms of historical setting, outcome measures, and overall results. By the middle of the nineteenth century, as the world's only fully industrialized nation, British output represented just under half the total of the world's industrial capacity. The Second Industrial Revolution in particular was an important episode in history that can provide important insights into the drivers of economic opportunities. We show that the railroad led to important transitions not only out of farming but into commercial and industrial activities. Connectivity has so far been measured as districts or provinces being connected. Our data allows us to measure connectivity at the individual level as the proximity between the address of residence and the nearest train station. This is especially important given that individuals can cross boundaries to get access to the railroad network. We show that even small distances away from the railroad network can have strong influence on social mobility patterns.

The rest of the paper is organized as follows. Section 2 paints the historical background of the railroad system in the nineteenth century England and Wales. It also describes our newly constructed datasets by linking several historical sources. Section 3 offers descriptives on intergenerational mobility including spatial clustering patterns. Section 4 presents the instrumental variable strategy we use to identify the causal effect of access to the railroad network and intergenerational mobility. Section 5 shows the significant role played by the railroad network on intergenerational mobility and its distributional consequences. We also investigate potential threats to our identification and the robustness of our results. Section 6 explores the mechanisms underlying our results. We finally summarize our findings and conclude in the last section.

2 Historical Background and Data

2.1 The Railroad Network

The British Industrial Revolution marked a period of development with profound social, economic and political change. Treiman (1970) suggests that industrialization involved the decline in the proportion of agricultural workers, created of a wider variety of occupations, generated more advantaged jobs and more educated workers, strengthened relationship between education and job, and weakened relationships between fathers and sons' job. The development of the railroad was an important driver of this transition.

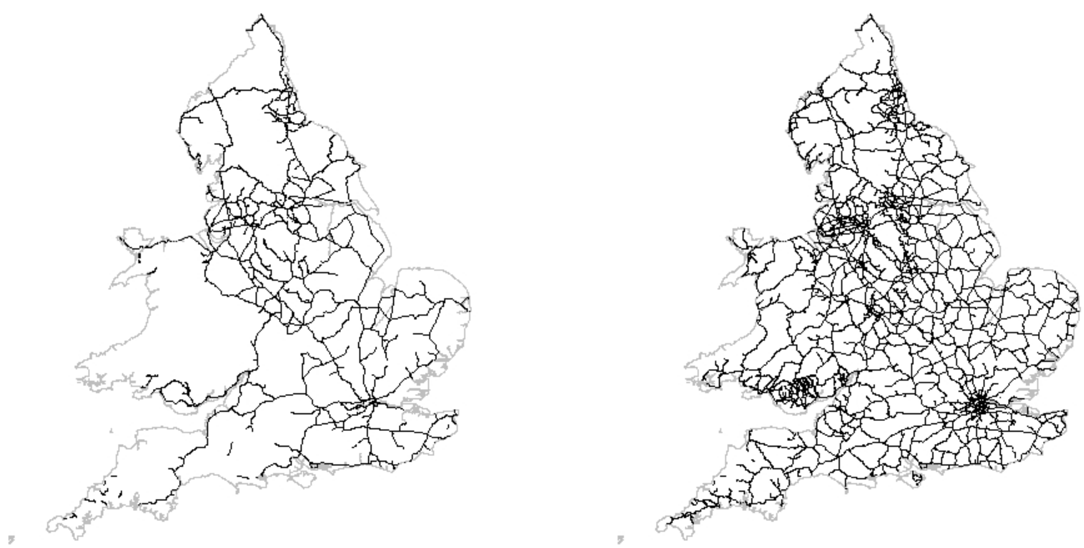
Britain was a pioneer in railroad technology and construction with inventors like Richard Trevithick and George Stephenson. The first steam-powered rail line was opened in 1825 between Stockton and Darlington in the northern coal mining region. By 1914 Britain had about 32,000 km of track. There was never a nationwide plan to develop a logical network of railroads. The railroad system was promoted by commercial interest and constructed entirely by private enterprises. Although the government initially took a laissez-faire approach, it was necessary to obtain an Act of Parliament to build a new railroad.

Railroads gave a great stimulus to local industries by enlarging the range of traffic that could be transported such as perishable goods, and reducing the freight costs of heavy materials such as coal and minerals. They were superior to canals as canals had been superior to roads for the carriage of heavy freight. The cost of canal carriage was 15 shillings a ton, whereas by rail it was 10 shillings a ton. By 1911, railroads conveyed c. 520 million tons of goods while canals only carried c. 40 million tons (Bagwell, 1974). Railways also facilitated the formation of an international inter-modal transport system by connecting major ports. The first example is the Liverpool-Manchester rail, which opened in 1830. It was the world's first public railroad to haul both passengers and freight. In particular it handled imports of raw cotton and exports of finished cotton goods by linking the Atlantic port of Liverpool to the textile centre of Manchester. The Newcastle and Carlisle Railway, was specifically built as a 'land bridge' to convey Scandinavian timber imported through the East Coast port of Newcastle to Ireland.

While the railroad was built mainly with freight in mind, passenger revenues exceeded 50% (Gourvish, 1988). The railroad brought affordable travel to a large proportion of the working population at unprecedented speeds. The Liverpool to Manchester journey took four hours, and cost 10 shillings inside the coach and 5 shillings outside. By train however, the same journey took one and three-quarter hours, and cost 5 shillings inside and 3 shillings 6 pence outside in 1830. As a point of reference, 5 shillings was the equivalent to a full week's work as a handloom weaver in 1831 or a full day's work as a textile factory worker in 1833 (Baines, 2015; Gaskell, 1836).[3] After the passing of Gladstone's Railway Act in 1844, which made the provision of third-class accommodation on at least one train per day obligatory at a cost of no more than a penny per mile, third-class passenger traffic took off going from 40 million to more than 1,200 million journeys from 1851 to 1911 (Bagwell, 1974).

The railroad network of England, Wales and Scotland was digitized by the Cambridge Group for the History of Population and Social Structure Alvarez et al. (2017). We exploit the railroad lines and stations of 1851 and 1881 as shown in Figure 1.

Figure 1: Railroad Network, 1851-1881



(a) 1851 (b) 1881

[3]With an average speed of less than 9mph by stagecoach on turnpike roads in 1830, travel from London took 15 hours to Birmingham, more than 20 hours to Manchester, and more than 30 hours to Newcastle. By contrast, with an average speed of 40 mph on the railroad at the beginning of the twentieth century, the same journeys from London took three, five and seven hours respectively (Bogart, Shaw-Taylor and You, 2018).

2.2 Intergenerational Mobility

2.2.1 Linking Individuals Across Censuses

Our aim is to relate an individual's access to the railroad network and their intergenerational mobility. For this purpose, we combine several historical sources to create a new and unique dataset. We first use the full-population censuses of England and Wales in 1851, 1881 and 1911 developed by the I-CeM project (Schürer and Higgs, 2020). The data contains records for 88 million individuals, and contains a wider range of sociodemographic information (age, gender, place of birth, marital status, number of children, number of servants and family structure), the full address of residence (house number or name, name of street, avenue or road, civil parish and county of residence), and self-reported occupation.

To create a measure of intergenerational mobility, we link individuals across consecutive censuses (1851–1881 and 1881–1911) using the matching procedure presented by Abramitzky et al. (2019). The linking strategy relies on four variables that should not change over time: birth year, county and parish of birth, given name, and surname. We focus on men, as women may have changed their surname due to marriage. Records are only compared in the linking process if they have an exact match on parish of birth, the difference in birth year is no larger than two years, and the first and last names have a Jaro-Winkler distance no larger than 0.1 (Jaro, 1989). Individuals are then matched across censuses if there is a unique match or the second best match is far enough, and there is no other person with a similar name within each census. As the censuses contain the household structure, we identify the sons or fathers of these linked men (see Appendix A.2 for further details). We impose the additional restriction that the distance in the surnames between fathers and sons is no larger than 0.12 in Jaro-Winkler to guarantee that the father-son pair are in fact from the same family. We also restrict sons to be between 40 and 52 years old and fathers to be between 20 and 65 years old in order to focus on men during their working years.

We link 980,848 father-son pairs, representing approximatively 43-50% of the population. As a point of comparison, match rates in other studies are between 7-42% (see Table A.3 for a comparison to other studies).[4] Section A.5 presents descriptives of the linked sample,

[4]The reason behind our higher match rate is the fact that, unlike historical US censuses where birthplace was listed at the state level, the UK censuses included birth parish. This much finer level increases the probability that a match will be unique. An additional advantage is the fact that we have a full census which reduces the probability of false positive, as pointed out by Bailey, Cole, Henderson and Massey (2020). Long (2005) also matches men English and Welsh census data from 1851 to 1911 and achieves a 15.2% to 33%. Their match rate is lower because they did not have access to the standardized birth parish variable recently constructed by I-CeM researchers, which addresses the issue of parishes with multiple and changing names.

showing that it is a representative sample of the full census in terms of spatial distribution and occupational ranking. In particular Table A.4 shows that the role of the railroad network in explaining the share of linked individuals is limited.[5]

2.2.2 Intergenerational Occupation Mobility

Linking individuals across censuses allows us to observe an individual's occupation as an adult (40-52 years old) and his father's occupation during his youth (10-22 years old). The 30 year interval allows the occupation information for both generations to be observed at a similar age. We measure intergenerational mobility through occupations as is commonly done in historical setting (Boberg-Fazlic, Sharp et al., 2013; Clark and Cummins, 2015; Ferrie, 2005; Long and Ferrie, 2013; Olivetti and Paserman, 2015). One of the advantage of using occupations is that they are more stable to transitory income shocks over the life cycle than income. Moreover, occupations can capture dimensions relevant to intergenerational mobility such as prestige in the community, autonomy in the workplace, and manual versus non-manual labor.

There are over 400 occupations reported in the census. We exploit both occupation ranking and occupation categories to measure intergenerational mobility. Occupations are ranked based on HISCAM (version 1.3.1 GB) which assigns a score to each occupation based on their position in the social stratification structure (Lambert, Zijdeman, Van Leeuwen, Maas and Prandy, 2013).[6] There are 359 unique HISCAM scores, and higher scores indicate a more advantageous position in society. Since we are interested in occupation mobility between father and son we employ a ranking that is constant over time. We define two indicator variables "upward mobility" and "downward mobility". The former (latter) switches from zero to one if the son's occupation has a higher (lower) score than the occupation of his father and the

[5]In addition to non-uniqueness, mortality and emigration are reasons why individuals are not matched. According to Woods and Hinde (1987), the probability of dying for males aged 10 and 29 was between 0.0248 and 0.0425 in 1838-54 and between 0.01 and 0.0263 in 1881-90. The life expectancy of a person age 10 was 47.05 in 1851 and 49 in 1881. There were approximately 27 and 84 emigrants per 10,000 between 1853 and 1910 (Snow, 1931). Among the 2,082,776 (3,346,899) individuals between the ages of 10 and 22 in 1851 (1881), we would not be able to link 2.7-5% (1.3-3.5%) because of death or emigration. In any case, survivor bias would only be a concern for our results if the proximity to the train station is somehow related to the survival probability.

[6]The HISCAM scale was derived using a method of "social interaction distance" analysis commonly used in sociology. Pairs of occupations linked by a social interactions such as marriage, friendship or parent-child relationship, are cross-tabulated and the frequency of occurrence is computed (e.g. how many bakers are friends of bakers, but also how many bakers are friends of butchers, secretaries...). Scores assigned to occupations represent the relative positions of those employed in each occupation, as revealed by the social interaction patterns. The HISCAM scores range from 28 to 99, with a mean of 50 and a standard deviation of 10.

difference in scores is higher than one standard deviation of the son's distribution.[7] We use the Historical International Standard Classification of Occupation (HISCO) to categorize occupations (Leeuwen, Maas and Miles, 2002). HISCO is not a class or status scheme but rather a classification by economic sector or workplace tasks. There are seven major groups: professional, managerial, clerical, sales, services, farm and laborer.[8] We also use alternative classifications to better capture class scheme, status and economic sectors (Woollard, 1998), skills (Van Leeuwen and Maas, 2011), and literacy requirements (Armstrong, 1972).

2.2.3 Geolocating Individuals

We geographically locate individuals down to the street level. For this we perform a string matching on address of residence (street name and parish) reported in the census and the digitized street points within each parish. The geo-referenced streets are based on the Great Britain addresses (GB1900) (Southall, Aucott, Fleet, Pert and Stoner, 2017), and the parish and county boundaries provided by the UK Data Service (Satchell, Kitson, Newton, Shaw-Taylor and Wrigley, 2017). Any measurement error in the location of individual can only occur within a parish. This high level of disaggregation allows us to measure individual access to the railroad network as the proximity between the place of residence and the nearest train station based on the shortest straight line. We are also able to capture geographic mobility from youth to adulthood.

3 Patterns of Intergenerational Mobility

Table 1 presents descriptive statistics of our sample. Sons and fathers are close in age. Sons grew up on average 3 km from a train station during their youth. 80% of sons do not follow their father's occupation, although both sons and fathers have on average an occupation rank of 50 and 49 respectively. Occupations ranked between 49 and 50 include a broad range of occupations such as farmer, laborers, professionals and services. 18% of sons experience

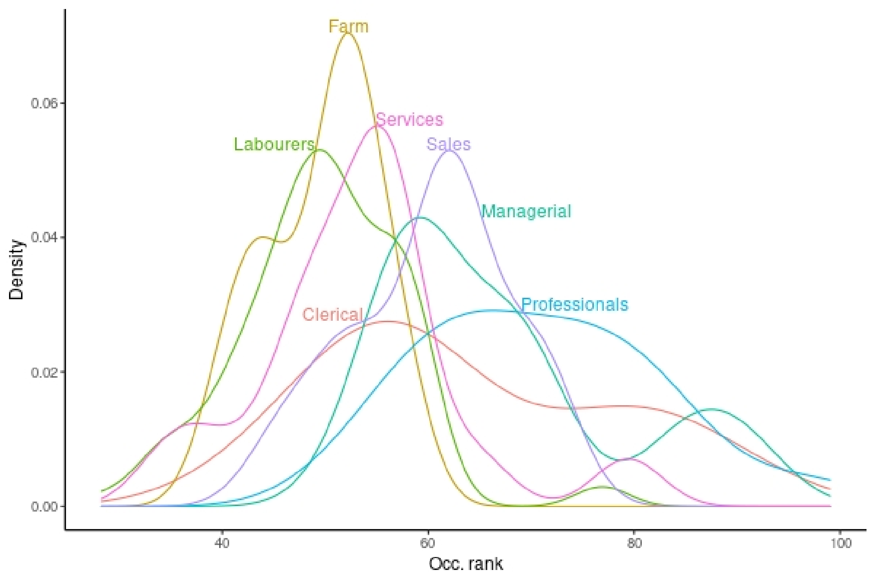
[7] More formally, let H^{son} be the HISCAM score of the son, with standard deviation $\sigma^{son} = \sqrt{Var(H^{son})}$. We define a son as upward mobile if $H^{son} > H^{father}$ and $|H^{son} - H^{father}| > \sigma^{son}$.

[8] "Professional" includes solicitors, clergy, accountants, high-wage merchants, "Managerial" include bankers, officers of commercial companies, manufacturers, other civil service officers and clerks, "Clerical" comprises commercial or business clerks, post officers and clerks, or messengers, "Sales" include grocers, commercial travellers, dealers, and insurance agents, "Services" include innkeepers, police, domestic servants, or hairdressers, "Agriculture" comprise of farm laborers and servants, "laborers" include for instance coal miners, carpenter, and painters.

upward mobility while 15% experience downward mobility. 31% of sons move away from the county they grew up in and move on average 100 km further away.

Figure 2 presents the distribution of the occupation ranking by occupation categories. We observe strong inequality between individuals at that time with very few individuals at the top of the distribution. We also see a clear ranking with professional occupations having on average a higher rank than agricultural occupations. Nevertheless, the ranking and category provide complementary information. Within each category, there is a range of ranks. For instance, within professional occupations monks have the highest rank while soldiers have the lowest rank.

Figure 2: Distribution of occupation ranking by occupation category, 1851-1911

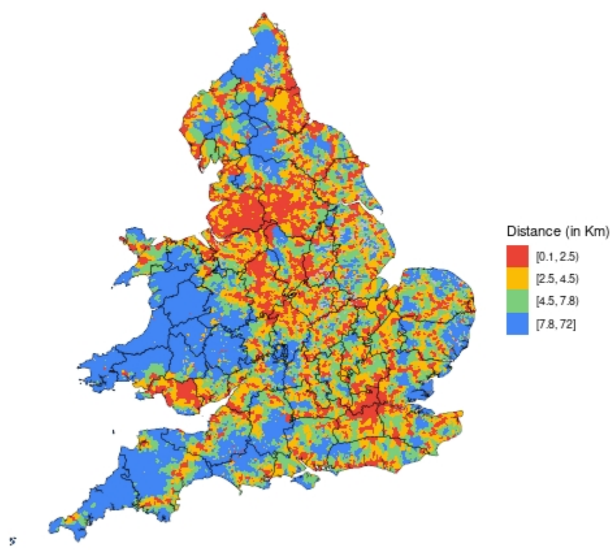


Note: This plots displays the density of HISCAM occupational rank by the HISCO occupation categories

The correlation between the occupation ranks of fathers and sons is 0.28. Table B.1 in the Appendix provides a cross-classification of sons and fathers' occupations. We distinguish between sons growing up within walking distance (i.e. 5 km) of a train station and those growing up further away. Regardless of connectedness, sons tended to follow their father's occupation as the larger percentage is found along the diagonal. Nevertheless, better connected sons experienced slightly greater mobility than those growing up further away from the train station. For instance, 36% of better connected sons whose fathers were farmers become lower skilled workers. In contrast, this share falls to 28% for sons growing up further away. Better connected sons whose fathers were in top (bottom) occupations were more (less) likely to stay in top occupations than sons who were less connected.[9]

A new feature of our dataset is the ability to geographically locate individuals. Figure 3 shows the average connectivity by parish. Most individuals lived within 5 to 10km to the nearest train station. Residents of Wales and Cornwall were the least connected to the railroad network. In contrast, residents of Manchester, Liverpool and Birmingham lived within 2.5km of the nearest train station. Figure 4 presents the share of sons who work in a different occupation than their father at the parish level and zooms in to the Liverpool to Manchester area. These figures reveals highly localized the spatial variation in intergenerational mobility.

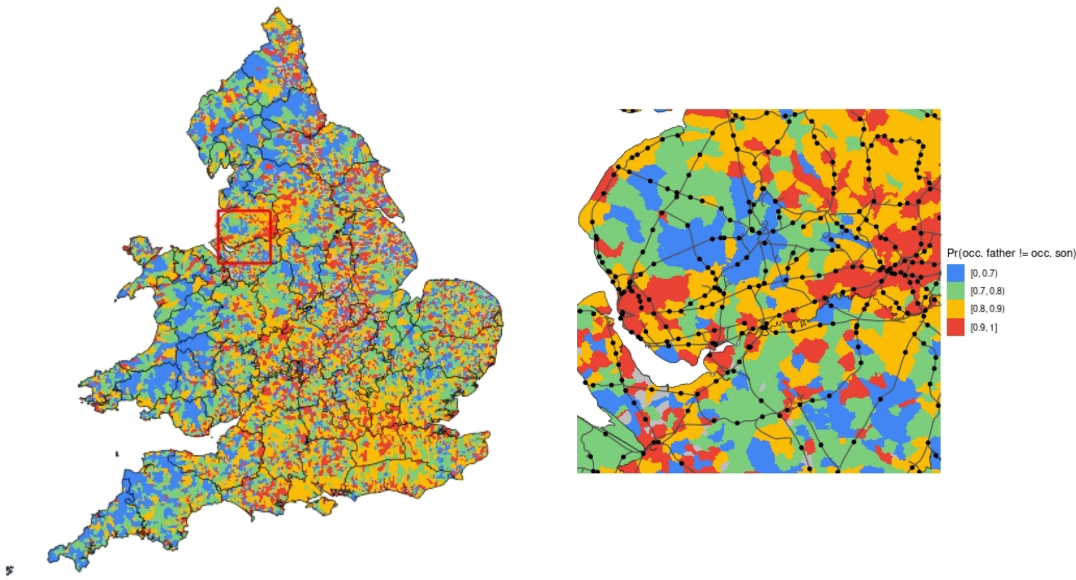
Figure 3: Avg. distance to the nearest train station (in km), 1851-1911



Note: This figure presents the average distance between place of residence during youth and the nearest train station by parish. Colors represents the quartiles. Borders in black represent counties.

[9]For an easier comparison to previous studies, we use the HISCLASS classification which capture the skills required in each occupation. We find 51-52% total mobility, XX% upward mobility and XX% downward mobility. This is similar to previous studies. Long (2013) measures the occupation intergenerational mobility for 1851-1881 and 1881-1901. He finds that the rate of total mobility is 48.3-50.1%, the rate of upward mobility is 26.8%, and the rate of downward mobility is 21.5-23.3%. Miles (1999) found a total mobility of 34.8% and upward mobility is 17.7% using a sample of marriage registries from 1859-1874.

Figure 4: Spatial intergenerational mobility pattern



Note: The figure to the left presents the share of sons working in a different occupation than their father at the parish level and the county borders. The figure to the right zooms in to the Liverpool-Manchester area with black lines and dots representing railway line and train stations. Colors represents the quartiles.

4 Empirical Strategy

To explore the role of the railroad network on intergenerational mobility, we estimate the following regression:

$$f(\text{Occ}^{son}_{i,c,t+1}, \text{Occ}^{father}_{i,c,t}) = \alpha \text{Proximity}_{i,c,t} + \beta X_{i,c,t} + \gamma_t + \rho_c + \epsilon_{i,c,t} \quad (1)$$

where i, c, and t index family (father-son pairs), county of residence, and census year when the father and son live together respectively. The dependent variable can take various forms: (1) an indicator variable equal to one if the son works in a different occupation category than his father, (2) the absolute difference in occupational ranks between the father and son, (3) an indicator variable equal to one if the son's occupational rank is larger than his father's and this difference is larger than one standard deviation of the son's distribution (i.e. upward mobility), (4) an indicator variable equal to one if the son's occupational rank is lower than his father's and this difference is larger than one standard deviation of the son's distribution (i.e. downward mobility).

We measure an individual's access to the railroad network using the standardized proximity between the place of residence and the nearest train station during their youth, $Proximity_{i,c,t}$. Specifically, we measure the straight line between the place where the sons grew up and their nearest train station. We winsorize the proximity at 10km, which is approximatively the top 5%. Our high spatial resolution allows us to be more precise than previous studies that measure access to the railroad network using an indicator variable for the presence of a train station or a railroad line in the district of residence. This is especially important given that individuals can cross district boundaries to access the railroad network. In alternative specifications, access to the railroad network is measured using indicators equal to one if the son grew up within 5, 10 and 15km of a train station or whether his parish of residence had a train station within its boundaries.

Finally, we include a vector of control variables $X_{i,c,t}$ which we discuss in Section 4.2. We also include census year γ_t and county ρ_c fixed effects. The former captures aggregate effects specific to sons in 1881 and those in 1911, which includes any overall improvement in labor opportunity due to the Industrial Revolution. The latter captures any time-invariant effects within a county such as the initial conditions including wealth, land suitability and local industries. Consequently, for two sons growing up in the same county during the same census year, the parameter α captures the effect of growing up one standard deviation (i.e. approximately 5km or one hour's walk) closer to the nearest train station on intergenerational mobility. There could be serial correlation in the error term $\epsilon_{i,c,t}$. We therefore cluster standard errors at the level of the parish of residence.

4.1 Dynamic Least Cost Railroad Network

Estimating Equation 1 using OLS would imply that, conditional on controls, year and county, the proximity to the railroads would have to be exogenous. Given the high cost and potential large benefits of infrastructure investments, the location of railroad lines and train stations was most likely correlated with the demand for trade, migration, local resources, and/or the cost of land. This raises the concern that connected locations were more likely to grow in the future, regardless of the railroad construction. It may also be the case that favorable labor market shocks happened to hit locations that were recently connected by the rail network, and this is what drives intergenerational mobility. In addition, individuals choose where to live and therefore their proximity to the nearest train station is endogenous. For instance, it may be that wealthier families, that experienced different mobility patterns, were more likely to live closer to town centres where the train station was generally located. If places or people with higher (lower) mobility potential were more likely to be connected to the railroad network, the OLS would overestimate (underestimate) the effect of being better connected.

To address this endogeneity issue, we use the "inconsequential place IV approach" (Alvarez et al., 2017; Banerjee et al., 2020; Chandra and Thompson, 2000; Faber, 2014; Lipscomb et al., 2013; Michaels, 2008). We construct a hypothetical railroad network showing how the railroad would have evolved had planners only considered geographic cost and ignored demand-side factors. We proceed in three steps. In the first step, we identify major towns in 1801 (Bennett, 2012).[10] By taking population at that time, we avoid any possible confounder related to population growth induced by the railroad. In the second step, we divide England and Wales into 50 $\times$ 50 grid cells. We construct least cost paths between all possible pairs of 1801 major towns imposing a cost to distance and altitude (Pope, 2017). The optimal path between two major towns is determined by minimizing the slope cost of all the cells the path crosses.[11] In a final step, we distinguish between rail lines that were likely to be constructed earlier than others. For this, we first compute the total slope cost of the actual 1851 network which serves as a budget. Rail lines with the highest edge betweenness are defined as "early" 1851 lines until the budget is exhausted.[12] The remaining least cost path network form the "late" 1881 projected lines.[13] The resulting dynamic least cost path network (DLCP) presented in Figure 5 is a function of the location of the 1801 population and geographic features of England and Wales. In the Appendix, we explore alternative instruments based solely on distance (not slope) or ignoring the dynamic feature.

While our $Proximity_{i,c,t}$ is defined as the proximity between the place of residence and the nearest train station, the instrument is defined as the proximity between the place of residence and nearest lines in the DLCP network. Therefore, the first stage equation is defined as:

$$Proximity_{i,c,t} = \delta(Proximity\ to\ DLCP)_{i,c,t} + \beta X_{i,c,t} + \gamma_t + \rho_c + \eta_{i,c,t} \tag{2}$$

where i, c and t index family, county and census year, respectively.

The instrument based on the DLCP railroad network isolates the portion of the variation in the expansion of the railroad network that is attributable to exogenous cost considerations.

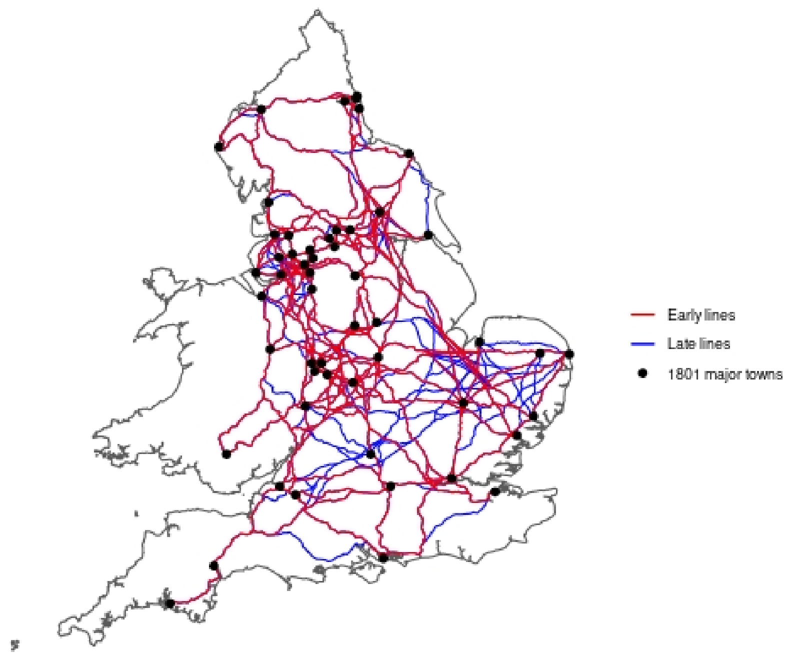
[10]Within all towns in 1801, we consider those in the top 10% of the population distribution (with at least 9,172 inhabitants) as a major town. There is a total of 53 major towns.

[11]The slope s has a cost of $1 + \left(\frac{s}{S}\right)^2$ for each cell crossing where S is a slope threshold that we set at the median slope of the observed network (Herzog, 2013).

[12]We rank rail lines (i.e. network edges) by decreasing order of edge betweenness, breaking ties using edge gravity, where this is defined as the product of populations the edge connects over square edge length. In doing so, we give priority to larger towns.

[13]When there are two lines within XX meter of one another, we merge them.

Figure 5: Dynamic Hypothetical Least Cost Path Network



Note: The black dots are the 1801 major towns. The lines represent the dynamic least cost path network. Red lines are the "early" 1851 lines and blue lines are the "late" 1881 lines.

In particular, the DLCP network is not based on local characteristics such as land value. Given that the instrument is defined as the proximity to nearest line in the DLCP network, it further decouples the location decision within towns. A family's location decision is unlikely to be correlated with the relative path to other town further away. This means that our inferences are based on individuals that are arbitrarily close to the railroad because they live on the least-cost path between end-nodes.

4.2 Identification Assumptions

The validity of the identification strategy depends on whether cost-side concerns can be fully separated from demand-side concerns within county and year. The exclusion restriction could be violated if locations along the least cost path between towns are correlated with economic characteristics due to history and/or sorting.

We used 1801 major towns as nodes in our DLCP network. This means that any individual residing between these nodes will mechanically be closer to important economic centres and will be more likely to lie on the DLCP than individuals living in towns further away. Proximity to major economic centres is likely to be correlated with town characteristics which also affect growth trajectories. This in turn will have a direct effect on the economic opportunities. We address this concern by including the distance to the closest 1801 town, and the 1801 population.[14]

The DLCP network is likely to follow pre-existing historical travel routes between cities. Any effects we attribute to being better connected to the network could in fact be due to the effects of being closer to other travel routes and not the railroad network. We control for the proximity to historical places of trade as proxied by ancient ports (Alvarez-Palau and Dunn, 2019), Roman Roads (McCormick, Huang, Zambotti and Lavash, 2013), and waterways (Satchell and Shaw-Taylor, 2018).

To the extent that the family characteristics such as initial wealth can determine both the place of residence and intergenerational mobility, the distance to the train station may be picking up family characteristics. We therefore control for household characteristics including the number of servants (a proxy of wealth generally used in historical settings), household size, and whether the father was born outside England and Wales.

In sum, the baseline identifying assumption is that residing along the DLCP network changes the economic outcomes from one generation to the next only through the railroad connection, conditional on the historical importance of towns, historical travel routes, household characteristics, county and year fixed effects.

5 Results

5.1 First Stage

In Table 2 we see a positive and statistically significant correlation between the proximity to the rail station and the proximity to the DLCP network. The instrument remains statistically significant and of similar magnitude with the inclusion of an increasingly comprehensive set of controls. The F-statistic on the first stage is large.

[14] The 1801 population is the population of the town and its surrounding area which is measured using the following equation: $\sum_{p \neq q} Pop_p / D_{p,q}$ where Pop_p is the standardized population of parish p and $D_{p,q}$ is the standardized distance between the centroids of parishes p and q.

5.2 Main Results

Our main results show that infrastructure in the form of access to the railroad network led to a break in the father-son occupational tie and significantly increase upward occupational mobility from one generation to the next. Table 3 presents the causal effect of being one standard deviation closer to the nearest train station on intergenerational mobility as estimated in Equation 1. The OLS results indicate that sons who grew up closer to a train station experienced significant change in occupation mobility. They were not only less tied to their father's occupation (row 1) but also moved further away from the occupation ranking of their father (row 2). Moreover, they experienced significant upward mobility (row 3) and little downward mobility (row 4) relative to their father. These effects become smaller in magnitude as we add more controls.

The results from our instrumental variable strategy paints a similar picture. Better connected sons experienced a significant break in ties to their father's occupation. The difference in occupational rank was also large and significant. This is largely due to an increase in upward mobility. As we include more control variables, the coefficients become smaller in magnitude. In our most restrictive specification we include all control variables in addition to county and census year fixed effects. This is our preferred specification for the remainder of the paper. Sons who grew up one standard deviation (approximately 5km or one hour's walk) closer to the train station were 11 percentage points more likely to work in a different occupation than their father. They were also 5 percentage points more likely to be upward mobile.[15]

The IV estimates identify a local average treatment effect among compliers. In our setup, this consists of individuals residing closer to the train station because their location was along a convenient route (i.e. close to the DLCP network) but would not have been so close otherwise.[16] Beyond providing a more accurate estimate of the effect of infrastructure on intergenerational mobility, the instrumental variable approach allows us to infer the direction and the magnitude of the selection due to non-random location of individuals with respect to the railroad network. The results from the OLS regressions underestimate the gains from connectivity, corroborating results from other studies.[17] This is consistent with

[15] We see these results as a linear approximation of a non-linear model for which we do not know the true thresholds. We explore non-linearities in section D.

[16] In 1881 (1911), 41% (75%) of sons grew up with a train station within their parish (roughly 2.5km to the nearest train station) and 33% (37%) grew up 2.5km from the nearest DLCP railroad line. In the robustness check, we compute the causal response weighting function.

[17] Other studies using using an "inconsequential place IV approach" to examine the effect of railroads have also found that the IV estimates were substantially larger than OLS estimates (e.g., Bogart et al., 2020;

the railroad locations targeting areas with limited intergenerational mobility. Historical evidence confirm that areas along the railroad and near train stations were negatively selected. Railroad companies wanted to connect large towns at the lowest cost. Consequently, they targeted low density places and cheaper land to save on the acquisition of land and the demolition of existing structures. Moreover, Acts of Parliament to build new railroads were often blocked by wealthy landowners with political power while local politicians generally lobbied to put stations in their constituency when it had low growth potential (Casson, 2009).[18] Alternatively, the OLS estimates could also be biased due to classical measurement error in the railroad access corrected by the IV estimate.

5.3 Change in Occupational Structure

The Industrial Revolution brought profound changes to the nature of work, and consequently the occupational structure of society. Before industrialization, the most significant economic activities were farming and artisan handicrafts. The coming of factory-based industry and other machinery set a shift with a decline in the proportion of agricultural workers and an increase in the prevalence of industrial and commercial activities. The railroad likely propelled this transition. Having established that connection to the railroad broke the link between fathers and sons' occupations and gave the opportunity to move upward in the occupational ranking, we next investigate its role in the transition between occupations. Specifically, we examine the probability that a son works in a specific occupation category:

$$\begin{aligned}\mathbb{I}[Occ^{son}_{i,c,t+1} \in cat] &= \alpha_1 Proximity_{i,c,t} + \alpha_2 \mathbb{I}[Occ^{father}_{i,c,t} \in cat] \\ &+ \alpha_3 Proximity_{i,c,t} \times \mathbb{I}[Occ^{father}_{i,c,t} \in cat] + \beta X_{i,c,t} + \gamma_t + \rho_c + \epsilon_{i,c,t}\end{aligned} \tag{3}$$

where cat refers to occupation categories. Just as in Equation 1, $Proximity_{i,c,t}$ is defined as the standardized proximity between the place of residence and the nearest train station when father and son lived together, and γ_t and ρ_c are the census year and county fixed effects. $X_{i,c,t}$ includes the complete set of control variables, namely the historical importance of town, historical travel routes and household characteristics. We include the effect of the father's occupation category an indicator variable $\mathbb{I}[Occ^{father}_{i,c,t} \in cat]$. We are particularly interested in α_3 which determines effect of proximity by the father's occupation category. We use a control function approach (Wooldridge, 2015) where in a first stage, we regress

Perez, 2017).

[18] According to Pollins (1952), 16% of the firms' costs came from buying land and 4% came from negotiation and lobbying expenditures in the parliament.

$Proximity_{i,c,t}$ on the proximity to the DLCP network and the complete set of controls as in Equation 2. In a second stage, we regress Equation 3 including the residual from the first stage as an additional control.[19]

Table 4 presents the transitions between occupations and reveals some interesting patterns. First, we see that access to the railroad network significantly affected the transition out of farming activities, and into professionals, clerical, sales and services activities regardless of the father's occupation. These results confirm that the railroad reinforced the change in occupational structure brought on by the Industrial Revolution. Second, we observe occupation upgrading thanks to the better access to the railroad network. For instance, better connected sons of salesmen were more likely be work in clerical and professional occupations. Third, we observe large variations in the access to the railroad network on the transition within occupations. For sons of managers, better access to the railroad meant that they were significantly likely to work in a service occupation and twice as likely to work in sales or professional occupations. Finally, there is also variation across occupations. Better access to the railroad network also increased the probability of becoming a laborer for sons of farmers. In contrast, it decreased the probability of becoming a laborer for sons whose fathers were working in the service occupations.

5.4 Distributional Effects

Occupational mobility may be driven by movements both from the bottom to the middle of the occupation ranking distribution and from the middle to the top of the occupation ranking distribution. These patterns have important implications for inequality patterns. To investigate distributional effects, we divide the occupational ranking into decile within the matched sample and census year, and we estimate equation 3.

Figure 6 presents the effect of proximity by the decile of the fathers' occupational ranking for each measure of occupational mobility. We see that the benefits from the access of the railroad network were not uniform across the occupational ranking distribution. The larger the father's occupation ranking, the larger the sons' benefits from being closer to the railroad network. For sons from the middle of the occupational ranking, better access to the railroad

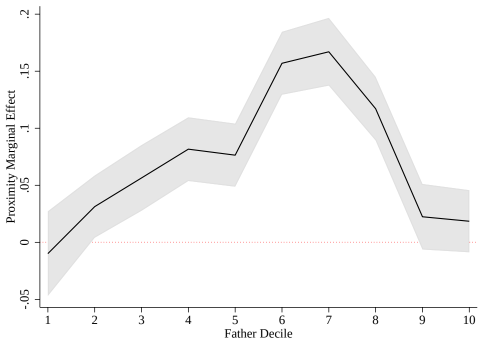
[19] The exogenous variation induced by the excluded DLCP network provides variation in the residuals obtained in the first stage and these residuals serve as the control function. By adding the control function, the endogeneity in the proximity to the railroad network becomes appropriately exogenous in a second-stage estimating equation. The instrumental variable and a control function approaches both produce the same baseline results (Wooldridge, 2015). We use a control function approach here in order to capture the interaction effects while keeping a strong first stage. We are taking into account the endogeneity in the fathers' occupations that is related to the proximity to the railroad network.

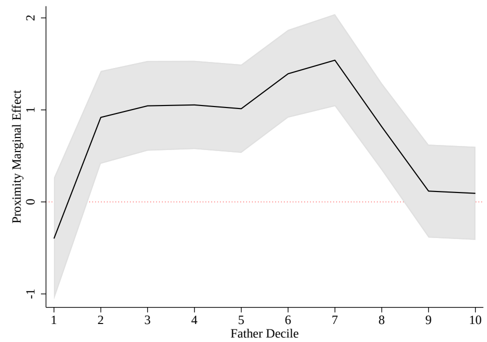
network meant that they were less tied to their father's occupation, and more likely to be upward mobile. In contrast, the benefits of the railroad network was limited for sons whose fathers were already at the top of the occupational rank. They remained tied to their father's occupation and did not move significantly in terms of occupational ranking.[20] In Table C.1 in the Appendix, we further investigate the differential effects by contrasting sons whose fathers were in white and blue-collar occupations.[21] Sons whose fathers was in a blue collar occupation experienced larger benefits from better access to the railroad network with higher occupational mobility and upward mobility than sons from white collar background.

[20]Figure C.1 in the Appendix displays the effect of proximity on the probability of working in a particular decile of the occupational ranking for sons and fathers separately. We see that for sons access to the railroad increased the probability of being at both ends of the HISCAM distribution while significantly decreasing the probability of being in the middle of the distribution. This middle of the occupation distribution is largely comprised of agricultural activities. The effect of connectivity to the railroad network on the fathers' HISCAM show a similar pattern. However, the negative effect close the 75th percentile is much more pronounced.

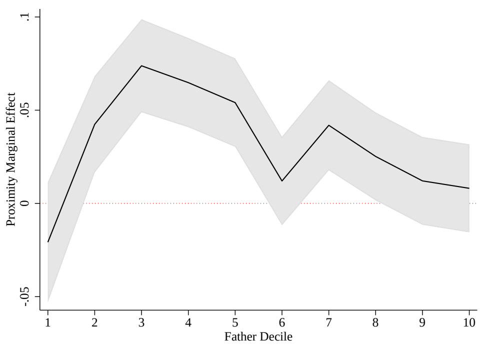
[21]White collar occupations is defined as HISCO between 0 and 5 and blue collar occupation as HISCO between 6 and 9.

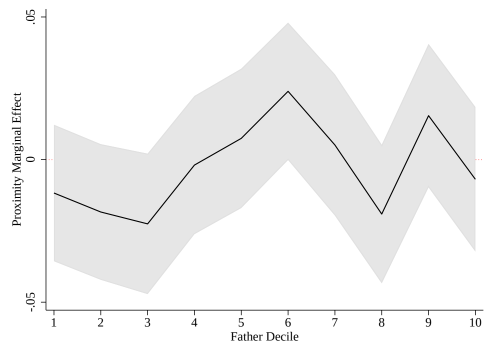
Figure 6: Distributional Consequences





(a) Occ. cat^{son} $\neq$ Occ. cat^{father} (b) |Occ. rank^{son} - Occ. rank^{father}|





(c) Upward mobility (d) Downward mobility

Note: This figure displays the coefficients and the 95% confidence interval of the effect of $Proximity_{i,c,t}$ on each measure of intergenerational mobility by the decile of the fathers' occupational rank (i.e. α_1 and α_3 of Equation 3). We use the control function approach where we interact the proximity with fathers' occupational quantile.

5.5 Robustness Checks

We perform a number of robustness checks. In all cases the same baseline result emerges: increased access to the railroad network led to a break between father and sons occupational tie, and a significant increase in upward mobility. Detailed explanations and results can be found in the Appendix D.

First, we show our baseline results are robust to alternative measures of connectedness to the railroad network, measures of intergenerational mobility, empirical specifications, and alternative instrumental variables. In Figure D.1, we define connectedness as the proximity to the nearest railroad line, an indicator variable equal to one if the son grew up within 5, 10 and 15km of a train station and whether the parish had a train station within its boundaries. In Figure D.2, we define upward (downward) mobility as an indicator variable taking the value one if the son has a higher (lower) occupational ranking than his father and the difference is at least 0.5, 1.5 or 2 standard deviation. We consider a HISCAM time-varying socio-economic status, capturing the fact that the position of an occupation may vary over time especially with the transition to industrialization. In Table D.1, we remove occupations specific to the railroad such as train conductor or controller which would mechanically increase with the expansion of the railroad network. We finally examine an alternative specification including polynomials for the control variables, and parish fixed effects in Figure D.3. Figure D.4 presents the results using alternative instruments including a least cost path network where the cost is based solely on distance and not the slope of the terrain, and a static least cost path network.

Second, we explore potential measurement errors in the geolocation of the place of residence and the linking procedure used to create father-son pairs. In Table D.2, we locate individual on the parish centroid instead of their address. In Table D.3, we control for the individual probability of being linked across censuses based on the proportion of linked individuals within county-of-birth, census-year and name-frequency. We see that the baseline results remain and the effects are similar in magnitude.

Third, in the presence of continuous, endogenous, and heterogeneous treatment effects, our linear IV estimate identifies a weighted average of causal responses (Angrist and Imbens, 1995). To understand how each observation contributes to our IV estimate, we compute the causal response weighting function following the decomposition proposed by Løken, Mogstad and Wiswall (2012). Figure D.5 shows weights for each level of proximity to the nearest train station. Figures D.6 explore possible non-linear effects.

Finally, we show that our results are robust to different subsamples: removing individuals living in 1801 major town (Table D.6), census year (Table D.4), county (Figure D.7), rural/urban divide (Table D.5), age of fathers and sons (Tables D.7 and D.8), birth order (Table D.9), natives/foreigners (Table D.10), locals/outsiders (Table D.11) or farming occupation (Table D.12).

6 Mechanisms

The previous section presented the causal evidence that the railroad network led to an increase in intergenerational occupation mobility. We now investigate how the railroad broke the link between father and son's occupational tie. Did better connectivity facilitate spatial mobility thereby increasing labor opportunities? Or did it improve local labor market prospects?

We decompose the effect access to the railroad network ($Train$) on intergenerational mobility (IM) between sons who move away from the county where they grew up ($Mover$) and those who stayed locally ($Stayer$).[22] Taking the total derivative with respect to train, we obtain:

$$\begin{aligned}\Delta \Pr(IM|Train) = & \underbrace{\Delta \Pr(IM|Stayer, Train)}_{\text{change in local opportunities brought by the train}} \\ & + \underbrace{[\Delta \Pr(IM|Mover, Train) - \Delta \Pr(IM|Stayer, Train)]}_{\text{change in the returns to spatial mobility induced by the train}} \times \underbrace{\Pr(Mover|Train)}_{\text{baseline spatial mobility}} \\ & + \underbrace{[\Pr(IM|Mover, Train) - \Pr(IM|Stayer, Train)]}_{\text{baseline returns to spatial mobility}} \underbrace{\times \Delta \Pr(Mover|Train)}_{\text{change in the spatial mobility thanks to the train}}\end{aligned} \tag{4}$$

The railroad therefore affects intergenerational mobility through three channels: (1) changes in local opportunities, (2) changes in the returns to spatial mobility, and (3) easing spatial mobility. In the following section, we estimate each component to understand the relative importance of each channel.

6.1 Returns to Spatial Mobility

The railroad network could have changed the relative benefit of moving by, for instance, bringing knowledge about job opportunities located further away. Measuring the return to spatial mobility is challenging given the selection issue. A naive comparison of sons who decided to move and those who decided to stay ignores the endogeneity in the decision to move. Movers may have earned more than stayers because bright and ambitious sons earn more regardless of their access to the railroad network but are also most likely to move. Following Abramitzky et al. (2012), we focus on brothers who grew up in the same household (Tables B.2 and C.2 in the Appendix presents the sample of brothers and their intergenerational mobility patterns). By comparing the outcome of sons who decided to move to their brothers who stayed, the estimate eliminates the component of the selection into migration that is shared between brothers such as financial constraints or unobserved ability. We therefore estimate the following equation

$$f(Occ^{son}_{i,t+1}, Occ^{father}_{i,t}) = \tau Mover^{son}_{i,t+1} + \lambda Proximity_{i,t} \times Mover^{son}_{i,t+1} + \psi_i + \epsilon_{i,t} \tag{5}$$

[22] $\Pr(IM|Train) = \Pr(IM|Stayer, Train) \times \Pr(Stayer|Train) + \Pr(IM|Mover, Train) \times \Pr(Mover|Train)$.

where i and t index family and census year when sons and fathers lived together respectively. The dependent variable takes the same four measures of intergenerational mobility as previously. The variable $Mover^{son}_{i,t+1}$ is an indicator variable equal to one if a son moves away from the county where he grew up and $Promixity_{i,t}$ is the proximity between the place of residence and the nearest train station. The family fixed effect ψ_i takes into account all within-family characteristics mentioned above. The coefficient τ represents the change in baseline returns to spatial mobility while λ estimates the change in the returns to spatial mobility from being better connected to the railroad network. We instrument the interaction between proximity and spatial mobility with the interaction of our DLCP instrument and $Mover^{son}_{i,t+1}$.

In Table 5 we see that there is a significant and positive return to spatial mobility for all measures of intergenerational mobility. In other words, brothers who moved were less tied to their father's occupation. They moved in occupations that were higher or lower in the occupational ranking than their father. However, we observe a negative return to spatial mobility from the proximity to the railroad network. For the brothers who moved, being closer to the railroad network decreased the intergenerational mobility. That is, they were more likely to follow their father's occupation and stay in the same occupational rank, but were also less downward mobile.

6.2 Spatial Mobility

From 1841 to 1901 the rural areas of England and Wales lost more than 4 million people from internal migration, 3 million of whom moved to towns, at a rate of more than half a million per decade (Crouzet, 2013). Railroads played an important role in these spatial mobility patterns by dramatically reducing travel time and cost. Bogart et al. (2020) for instance find that having a railroad station in a locality by 1851 in England and Wales led to significantly higher population growth from 1851 to 1891. To explore the role of the railroad on spatial mobility, we look at the probability of sons moving away from the county where they grew up

$$\mathbb{I}[Mover^{son}_{i,c,t+1}] = \phi Proximity_{i,c,t} + \beta X_{i,c,t} + \gamma_t + \rho_c + \epsilon_{i,c,t} \tag{6}$$

where $Mover^{son}_{i,c,t+1}$ is an indicator that takes a value of 1 if a son resided in a different county from the one he grew up in. All independent variables are the same as in equation 1. We instrument $Proximity_{i,c,t}$ using the DLCP railroad network as in equation 2.

Table 6 shows that better access to the railroad network eased the spatial mobility of residents. Sons who grew up 5km closer to the train station were 15 percentage points more likely to move away from the county where they grew up. It is reasonable to ask whether a one-time migration cost, which may be small relative to the present value of a higher future income stream, will affect the decision to move away. Similarly to Morten and Oliveira (2014), we think of migration costs broadly to include both financial and utility costs of moving such as the costs related to being away from friends and family (e.g. return visits which are costly in terms of time and money) and the costs of not being able to consume the same types of goods as at home.[23]

6.3 Local Labor Opportunities

The railroad network could have improved local labor opportunities in various ways. First, labor opportunities now became "local" thanks to the railroad. It offered the possibility of commuting which created a separation of the home from the workplace (Heblich et al., 2020). Transport infrastructure also generate well-known agglomeration effects, in which the dense population of urban areas has an effect on the productivity of resources. We find that the railroad network promoted urbanization, confirming previous results by Alvarez et al. (2017) (column 1 of Table 7).

Second, railroad networks integrate local economies with external markets (Donaldson, 2018; Donaldson and Hornbeck, 2016). Falling transport costs reduced the cost of raw materials and expanded the size of markets, eroding monopoly power and compelling firms to raise productivity through division of labor. Railroads have spurred the rise of factories (Atack et al., 2020) and facilitated information flows and the adoption of new technologies (Agrawal, Galasso and Oettl, 2017; Andersson, Berger and Prawitz, 2021). This created a class of wealthy entrepreneurs and comfortable middle class supported by workers (Hanlon, 2020). We find evidence that the railroad powered industrialization in Table 7. We see that the railroad network let to a significant increase in firm concentration (proxied by the number of chimneys in a parish (Heblich, Trew and Zylberberg, 2021)) and the number of entrepreneurs. Consequently, it altered the social structure of society, with a higher local occupational rank and inequality. To explore the role of the railroad network on the creation of new industries, we also look at transitions between occupations that grew or decline over

[23]Spatial mobility, especially for poor individuals, was limited by the Law of Settlement, which sanctioned the removal of unsettle poor who would be an economic burden to a parish. However, by 1864, the scope of Law of Settlement had been greatly attenuated (Feldman, 2003).

our sample period.[24] In Table C.3, we see that sons who grew up closer to the railroad network were 15% less likely to work in a declining occupation and 5% more likely to work in a growing occupation, regardless of their father's occupation.

Third, these new industries which created new job opportunities likely required new skills. It has been shown that the railroad lead to higher school enrolment and increase skill premia in the local labor market (Adukia, Asher and Novosad, 2020; Atack, Margo and Perlman, 2012; Chaudhary and Fenske, 2020; Michaels, 2008). Tables C.5 and C.4 point to the fact that occupational upgrading in term of categories and ranking is partly explained by such educational and skill investments. Better connected sons were 8 percentage points more likely to be literate and 4 percentage points more likely to work in a high-skilled occupation.[25]

To examine the relative size of the three channels at work, we decompose estimate the magnitude of the effect of the railroad into the three channels at work. The majority of intergenerational mobility induced by the railroad network is driven by changes in the local labor market opportunities. In particular, local opportunities account for roughly 97% of the upward mobility and 78% of the occupation mobility (see Table C.6 in the Appendix).[26]

7 Conclusion

The long-run implications of infrastructure improvements for economic opportunities of individuals is of interest both for historical reasons and also because they are related to current debates on institutional change. Can transport infrastructure break the link between parents and their children's economic outcomes? This paper is the first to estimate the causal effect of the railroad network on intergenerational mobility in nineteenth century England and Wales.

Understanding the effect of infrastructure on intergenerational mobility is empirically challenging due to data availability and non-random placement of infrastructure. We create

[24] Growing/declining occupations are those that are at the top/bottom 25% of the change in the share of occupation between 1851 and 1911. Table B.3 in the Appendix presents examples of occupations with the highest and lowest growth.

[25] Skill level is defined as an indicator variable equal to one if the HISCLASS occupational ranking is "manager", "skilled worker" or "lower skilled". Literacy is based on Armstrong (1972)'s measure of the literacy requirement for each occupation.

[26] This remains a decomposition exercise. Although we address the endogeneity issue in the decision to move, we do not take into account the endogenous destination location. The destination location is likely correlated with the individual's skill set and the complementarities in the labor opportunity. Therefore the relative benefit of moving should take into account the specific place of origin and destination. Moreover, there may be general equilibrium spatial spillover effects where the construction of a new line affects not only the local area but also the other areas.

a new dataset which allows us to observe the occupation of father-son pair between 1851 and 1911 and geographically locate them down to the street level. This new level of disaggregation allows us to measure access the railroad network based on the proximity to the nearest train station. To address the endogenous access to the railroad, we create a dynamic least-cost railroad network. This allows us to isolate the portion of the variation that is attributable to exogenous cost considerations and use it as an instrument.

We find that railroads led to significant changes on intergenerational mobility patterns. Our results highlight that the effects of access to transport infrastructure is highly localized, heterogenous and long-lasting. Sons who grew up one standard deviation (approximately 5km or one hour's walk) closer to the nearest train station were 11 percentage points more likely to work in a different occupation as their father. They were also 5 percentage points more likely to be upward mobile. These effects are not only driven by significant transitions out of farming activities, but also transitions into industrial and commercial activities. The benefits of the railroad access was not uniform, particularly benefitting sons from the middle of occupational ranking backgrounds.

When decomposing the intergenerational mobility into the various channels at work, we find that the majority of the effect is driven by changes in the labor opportunities brought to town by the railroad or becoming feasible by commuting. We find that the railroad network allowed people to flock to cities, industries to expand, and a new class of wealthy entrepreneurs to form. Consequently, the railroad altered the social structure with higher local occupational ranks and inequality.

References

Abramitzky, R., Boustan, L.P., Eriksson, K., 2012. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. American Economic Review 102, 1832–56.

Abramitzky, R., Boustan, L.P., Eriksson, K., 2014. A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. Journal of Political Economy 122, 467–506.

Abramitzky, R., Mill, R., Pérez, S., 2019. Linking individuals across historical sources: A fully automated approach. Historical Methods: A Journal of Quantitative and Interdisciplinary History , 1–18.

Adukia, A., Asher, S., Novosad, P., 2020. Educational investment responses to economic opportunity: evidence from indian road construction. American Economic Journal: Applied Economics 12, 348–76.

Agrawal, A., Galasso, A., Oettl, A., 2017. Roads and innovation. Review of Economics and Statistics 99, 417–434.

Alesina, A., Hohmann, S., Michalopoulos, S., Papaioannou, E., 2021. Intergenerational mobility in Africa. Econometrica 89, 1–35.

Alvarez, E., Bogart, D., Satchell, M., Shaw-Taylor, L., You, X., 2017. Railways and growth: Evidence from 19th Century England and Wales. mimeo .

Alvarez-Palau, E.J., Dunn, O., 2019. Database of historic ports and coastal sailing routes in England and Wales. Data in brief 25, 104188.

Andersson, D., Berger, T., Prawitz, E., 2021. Making a market: Infrastructure, integration, and the rise of innovation. Review of Economics and Statistics , 1–44.

Angrist, J.D., Imbens, G.W., 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. Journal of the American Statistical Association 90, 431–442.

Armstrong, W., 1972. The use of information about occupation. in Nineteenth Century Society, ed. E.A. Wrigley (Cambridge: Cambridge University Press, 1972), p. 255-81 .

Atack, J., Haines, M., Margo, R.A., 2020. 7. railroads and the rise of the factory: Evidence for the united states, 1850–1870, in: Economic Evolution and Revolution in Historical Time. Stanford University Press, pp. 162–179.

Atack, J., Margo, R., Perlman, E., 2012. The impact of railroads on school enrollment in 19th Century America. Unpublished Working Paper, Boston University .

Bagwell, P.S., 1974. The transport revolution from 1770. London: Batsford.

Bailey, M.J., Cole, C., Henderson, M., Massey, C., 2020. How well do automated linking methods perform? Lessons from US historical data. Journal of Economic Literature 4, 997–1044.

Baines, E., 2015. History of the cotton manufacture in Great Britain. Cambridge University Press.

Baker, W.J., 1971. Railways and economic growth in england and wales 1840-1870. The Journal of Economic History 31, 718–719.

Banerjee, A., Duflo, E., Qian, N., 2020. On the road: Access to transport infrastructure and economic growth in China. Journal of Development Economics 145, 102442.

Barone, G., Mocetti, S., 2021. Intergenerational mobility in the very long run: Florence 1427–2011. The Review of Economic Studies 88, 1863–1891.

Baum-Snow, N., 2007. Did highways cause suburbanization? The Quarterly Journal of Economics 122, 775–805.

Bennett, R.J., 2012. Urban population database, 1801-1911. [data collec- tion] Robson, B., University of Manchester, Department of Geography. SN: 7154, http://doi.org/10.5255/UKDA-SN-7154-1 .

Björklund, A., Jäntti, M., 1997. Intergenerational income mobility in sweden compared to the united states. The American Economic Review 87, 1009–1018.

Boberg-Fazlic, N., Sharp, P., et al., 2013. North and south: Social mobility and welfare spending in preindustrial England. European Historical Economics Society (EHES) Working Paper 37.

Bogart, D., Shaw-Taylor, L., You, X., 2018. The development of the railway network in britain 1825-1911. Transport, Urbanization and Economic Development in England and Wales C , 1670–1911.

Bogart, D., Xuesheng, Y., Alvarez, E., Satchell, M., Shaw-Taylor, L., 2020. Railways, population divergence, and structural change in 19th century England and Wales. mimeo .

Casson, M., 2009. The world's first railway system: Enterprise, competition, and regulation on the railway network in Victorian Britain. Oxford University Press.

Chandra, A., Thompson, E., 2000. Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. Regional science and urban economics 30, 457–490.

Chatterjee, S., Turnovsky, S.J., 2012. Infrastructure and inequality. European Economic Review 56, 1730–1745.

Chaudhary, L., Fenske, J., 2020. Did railways affect literacy? evidence from India. Warwick Economics Research Papers .

Chetty, R., Hendren, N., 2013. The economic impact of tax expenditures: Evidence from spatial variation across the US .

Chetty, R., Hendren, N., 2018. The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. The Quarterly Journal of Economics 133, 1107–1162.

Chetty, R., Hendren, N., Kline, P., Saez, E., 2014. Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. The Quarterly Journal of Economics 129, 1553–1623.

Clark, G., Cummins, N., 2015. Intergenerational wealth mobility in England, 1858–2012: Surnames and social mobility. The Economic Journal 125, 61–85.

Crouzet, F., 2013. The Victorian Economy. Routledge.

Donaldson, D., 2018. Railroads of the Raj: Estimating the impact of transportation infrastructure. American Economic Review 108, 899–934.

Donaldson, D., Hornbeck, R., 2016. Railroads and American economic growth: A market access approach. The Quarterly Journal of Economics 131, 799–858.

Duranton, G., Turner, M.A., 2012. Urban growth and transportation. Review of Economic Studies 79, 1407–1440.

Faber, B., 2014. Trade integration, market size, and industrialization: evidence from China's national trunk highway system. Review of Economic Studies 81, 1046–1070.

Feldman, D., 2003. Migrants, immigrants and welfare from the old poor law to the welfare state. Transactions of the Royal Historical Society , 79–104.

Ferrie, J.P., 2005. History lessons: The end of American exceptionalism? Mobility in the United States since 1850. The Journal of Economic Perspectives 19, 199–215.

Gaskell, P., 1836. Artisans and machinery: The moral and physical condition of the manufacturing population considered with reference to mechanical substitutes for human labour. JW Parker.

Gourvish, T.R., 1988. Railways 1830-70: the formative years. Transport in Victorian Britain , 82.

Grawe, N.D., 2006. Lifecycle bias in estimates of intergenerational earnings persistence. Labour economics 13, 551–570.

Güell, M., Rodríguez Mora, J.V., Telmer, C.I., 2015. The informational content of surnames, the evolution of intergenerational mobility, and assortative mating. The Review of Economic Studies 82, 693–735.

Guerra, J.A., Mohnen, M., 2020. Multinomial choice with social interactions: Occupations in Victorian London. The Review of Economics and Statistics (forthcoming).

Hanlon, W.W., 2020. The rise of the engineer: Inventing the professional inventor during the industrial revolution. mimeo .

Heblich, S., Redding, S.J., Sturm, D.M., 2020. The making of the modern metropolis: evidence from london. The Quarterly Journal of Economics 135, 2059–2133.

Heblich, S., Trew, A., Zylberberg, Y., 2021. East-side story: Historical pollution and persistent neighborhood sorting. Journal of Political Economy 129, 1508–1552.

Herzog, I., 2013. The potential and limits of optimal path analysis. Computational approaches to archaeological spaces 186.

Jaro, M.A., 1989. Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. Journal of the American Statistical Association 84, 414–420.

Lambert, P.S., Zijdeman, R.L., Van Leeuwen, M.H., Maas, I., Prandy, K., 2013. The construction of HISCAM: A stratification scale based on social interactions for historical comparative research. Historical Methods: A Journal of Quantitative and Interdisciplinary History 46, 77–89.

Leeuwen, M., Maas, I., Miles, A., 2002. HISCO: Historical international standard classification of occupations. Leuven: Leuven University Press.

Levine, P.B., Zimmerman, D.J., et al., 1996. The Intergenerational Correlation in AFDC Participation: Welfare Trap or Poverty Trap? University of Wisconsin-Madison, Institute for Research on Poverty.

Lipscomb, M., Mobarak, A.M., Barham, T., 2013. Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil. American Economic Journal: Applied Economics 5, 200–231.

Løken, K.V., Mogstad, M., Wiswall, M., 2012. What linear estimators miss: The effects of family income on child outcomes. American Economic Journal: Applied Economics 4, 1–35.

Long, J., 2005. Rural-urban migration and socioeconomic mobility in Victorian Britain. The Journal of Economic History 65, 1–35.

Long, J., 2013. The surprising social mobility of Victorian Britain. European Review of Economic History 17, 1–23.

Long, J., Ferrie, J., 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. The American Economic Review 103, 1109–1137.

Machin, S., 2007. Education expansion and intergenerational mobility in Britain. Schools and the Equal Opportunity Problem. The MIT Press: Cambridge 7, 1–63.

McCormick, M., Huang, G., Zambotti, G., Lavash, J., 2013. Roman road network (version 2008). DARMC Scholarly Data Series, Data Contribution Series 5.

Michaels, G., 2008. The effect of trade on the demand for skill: Evidence from the interstate highway system. The Review of Economics and Statistics 90, 683–701.

Miles, A., 1999. Social mobility in nineteenth-and early twentieth-century England. Springer.

Milner, B., 2020. The impact of state-provided education: Evidence from the 1870 Education Act. mimeo .

Mitch, D., 1992. The rise of popular literacy in Victorian England: the influence of private choice and public policy. University of Pennsylvania Press.

Morten, M., Oliveira, J., 2014. Migration, roads and labor market integration: Evidence from a planned capital city. Unpublished manuscript .

Olivetti, C., Paserman, M.D., 2015. In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850-1940. American Economic Review 105, 2695–2724.

Perez, S., 2017. Railroads and the rural to urban transition: Evidence from 19th Century Argentina. mimeo .

Piketty, T., 2000. Theories of persistent inequality and intergenerational mobility. Handbook of income distribution 1, 429–476.

Pollins, H., 1952. A note on railway constructional costs 1825-1850. Economica 19, 395–407.

Pope, A., 2017. SRTM Slope DEM for Great Britain.

Prandy, K., Bottero, W., 2000. Social reproduction and mobility in Britain and Ireland in the nineteenth and early twentieth centuries. Sociology 34, 265–281.

Redding, S.J., Turner, M.A., 2015. Transportation costs and the spatial organization of economic activity, in: Handbook of regional and urban economics. Elsevier. volume 5, pp. 1339–1398.

Rostow, W.W., 1959. The stages of economic growth. The Economic History Review 12, 1–16.

Sanderson, E., Windmeijer, F., 2016. A weak instrument F-test in linear IV models with multiple endogenous variables. Journal of Econometrics 190, 212–221.

Satchell, A., Kitson, P., Newton, G., Shaw-Taylor, L., Wrigley, E., 2017. 1851 England and Wales census parishes, townships and places. [data collection]. UK Data Service. SN: 852816, http://doi.org/10.5255/UKDA-SN-852232 .

Satchell, M., Shaw-Taylor, L., 2018. 1820 england and wales navigable waterways. [Data Collection]. Colchester, Essex: UK Data Archive. 10.5255/UKDA-SN-852997 .

Schürer, K., Higgs, E., 2020. Integrated census microdata (I-CeM); 1851-1911. UK Data Service. [data collection]. DOI: 10.5255/UKDA-SN-7481-2 Colchester, Essex: UK Data Archive [distributor], April 2014. SN: 7481.

Sequeira, S., Nunn, N., Qian, N., 2020. Immigrants and the making of america. The Review of Economic Studies 87, 382–419.

Snow, C., 1931. Emigration from Great Britain, in: International Migrations, Volume II: Interpretations. NBER, pp. 237–260.

Southall, H., Aucott, P., Fleet, C., Pert, T., Stoner, M., 2017. GB1900: Engaging the public in very large scale gazetteer construction from the ordnance survey county series 1: 10,560 mapping of Great Britain. Journal of Map & Geography Libraries 13, 7–28.

Treiman, D.J., 1970. Industrialization and social stratification. Sociological inquiry 40, 207–234.

Van Leeuwen, M.H., Maas, I., 2011. HISCLASS: A historical international social class scheme. Universitaire Pers Leuven.

Winkler, W.E., 1999. The state of record linkage and current research problems, in: Statistical Research Division, US Census Bureau, Citeseer.

Woods, R., Hinde, P.A., 1987. Mortality in Victorian England: Models and patterns. Journal of Interdisciplinary History , 27–54.

Wooldridge, J.M., 2015. Control function methods in applied econometrics. Journal of Human Resources 50, 420–445.

Woollard, M., 1998. The classification of occupations in the 1881 census of England and Wales. History and Computing 10, 17–36.

Xie, Y., Killewald, A., 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment. American Economic Review 103, 2003–20.

Table 1: Descriptive Table

	Mean	St.Dev.	Min.	Median	Max.
			A.SONS		
Age	44.64	3.51	40.00	44.00	52.00
Foreign-born	0.02	0.15	0	0	1.00
Urban resident	0.22	0.42	0	0	1.00
Literate	0.38	0.49	0	0	1.00
Occ. rank	50.09	10.11	28.28	50.81	99.00
Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}	0.79	0.40	0	1.00	1.00
$\vert$Occ. rank ^{son} - Occ. $\text{rank}^{father}\vert$	8.02	8.38	0	5.89	70.72
Upward mobility	0.18	0.39	0	0	1.00
Downward mobility	0.15	0.36	0	0	1.00
County mover	0.31	0.46	0	0	1.00
Dist. moved $\vert$ county mover (in km)	100.32	97.30	0.04	70.59	663.55
Dist. to nearest train station (in km)	3.28	5.45	0	1.50	83.53
			B.FATHERS		
Age	46.67	7.61	20.00	46.00	65.00
Foreign-born	0.05	0.22	0	0	1.00
Urban resident	0.17	0.37	0	0	1.00
Household size	6.75	2.15	0	7.00	19.00
Number of servants	0.16	0.62	0	0	15.00
Literate	0.33	0.47	0	0	1.00
Occ. rank	49.41	9.11	28.28	50.95	99.00
			C.COUNTY		
Number of father-son pairs	8,917	13,119	88	4,766	88,642
Area (km^2)	2,739	1,598	2	2,213	7,136
Population	563,116	805,366	18,869	325,073	4,664,121
Avg. occ. rank	49.93	1.71	45.33	50.37	53.52
Avg. dist. to train station (in km)	6.92	9.06	1.27	3.89	48.26

Note: The sample consists of 980,848 father-sons pairs living in 55 counties. Sons are 10-22 years old when their father's occupation is measured in 1851 or 1881, and 40-52 years old when their own occupation is measured in 1881 or 1911. The table provides descriptives for the sons as adult (panel A), fathers (panel B), and county (panel C).

Table 2: First stage regressions

	(1)	(2)	(3)
		Proximity$_{i,c,t}$	
Proximity to DLCP network$_{i,c,t}$	0.640	0.339	0.339
	(0.031)	(0.029)	(0.029)
F-Stat	414.450	135.729	135.680
Obs.		980,848	
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel routes	No	Yes	Yes
Household characteristics	No	No	Yes

Note: The dependent (independent) variable is the standardized proximity between the residence during youth and the nearest train station (the nearest railroad line from the DLCP network). All regressions include fixed effects for census year and county. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its population and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2 and 3), household characteristics including the number of servants, household size and whether the father is born outside England and Wales (column 3). Standard errors clustered at the parish level are reported in parentheses. F-stat reports the F-statistic from Sanderson and Windmeijer (2016).

Table 3: The effect of railroad connection on intergenerational mobility

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}	0.056	0.027	0.026	0.151	0.113	0.106
	(0.003)	(0.003)	(0.003)	(0.013)	(0.023)	(0.023)
$\lvert$Occ. rank^{son} - Occ. $\text{rank}^{father}\rvert$	1.050	0.696	0.683	2.034	1.419	1.313
	(0.044)	(0.042)	(0.042)	(0.164)	(0.301)	(0.295)
Upward Mobility	0.041	0.029	0.029	0.068	0.048	0.044
	(0.002)	(0.002)	(0.002)	(0.007)	(0.013)	(0.013)
Downward Mobility	0.014	0.007	0.007	0.038	0.029	0.028
	(0.001)	(0.001)	(0.001)	(0.005)	(0.011)	(0.011)
Obs.	980,846					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each cell represents the coefficient of the standardized proximity to the nearest train station (columns 1 to 3) and instrumented by the proximity to the DLCP railroad network (columns 4 to 6). The dependent variable is an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the occupational rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and county. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5 and 6), and household characteristics including the number of servants, household size and whether the father is born outside England and Wales (columns 3 and 6). Standard errors clustered at the parish level are reported in parentheses.

Table 4: The effect of rail connection by HISCO occupations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Father							
Son	Professional	Managerial	Clerical	Sales	Services	Farm	Labourer	Any
Professionals	-0.026	0.032	0.013	0.018	0.011	0.012	0.011	0.018
	(0.009)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.006)
Managerial	-0.005	0.006	-0.005	0.000	0.000	-0.003	-0.003	-0.003
	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Clerical	0.033	0.048	0.037	0.029	0.035	0.019	0.019	0.031
	(0.007)	(0.007)	(0.010)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
Sales	0.018	0.032	0.021	0.005	0.016	0.020	0.016	0.040
	(0.009)	(0.009)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
Services	-0.003	0.013	0.001	0.017	0.028	0.022	0.018	0.024
	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.007)	(0.008)	(0.008)
Farm	-0.015	-0.126	-0.030	-0.057	-0.052	-0.108	-0.052	-0.168
	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.031)
Labourers	-0.002	-0.005	-0.038	-0.013	-0.038	0.038	-0.009	0.058
	(0.019)	(0.019)	(0.020)	(0.018)	(0.018)	(0.017)	(0.018)	(0.026)
Obs.	980,848							

Note: Each entry represents the coefficient of the standardized proximity, instrumented by the proximity to the DLCP network in a control function approach. The dependent variable is an indicator equal to one if the son works in a specific HISCO occupation (rows). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier (column 8). We also report estimates by father occupation: professional (column 1, 21,508 obs.), managerial (column 2, 230,350 obs.), clerical (column 3, 580,193 obs.), sales (column 4, 15,528 obs.), services (column 5, 22,509 obs.), farm (column 6, 75,856 obs.), and labourer (column 7, 34,904 obs.). All regression include census year and county fixed effects. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway; the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish level are reported in parentheses.

Table 5: Returns to spatial mobility

	(1)	(2)	(3)	(4)
	Occ. cat^{son} $\neq$ Occ. cat^{father}		$\lvert$Occ.$rank^{son}$ - Occ. $rank^{father}\rvert$	
$Mover^{son}_{i,t+1}$	0.103	0.041	1.436	0.790
	(0.004)	(0.009)	(0.068)	(0.168)
$Mover^{son}_{i,t+1}$ * $Proximity_{i,t}$		-0.123		-1.301
		(0.016)		(0.316)
	(5)	(6)	(7)	(8)
	Upward Mobility		Downward Mobility	
$Mover^{son}_{i,t+1}$	0.040	0.030	0.016	0.001
	(0.003)	(0.007)	(0.003)	(0.007)
$Mover^{son}_{i,t+1}$ * $Proximity_{i,t}$		-0.019		-0.029
		(0.014)		(0.013)
F-Stat		311.689		311.689
Obs.	342,715			

Note: The dependent variable is an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1 and 2), the absolute value of the difference in the occupation rank between sons and fathers (columns 3 and 4), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and the difference is greater than one standard deviation (columns 5 and 6 / 7 and 8). $Mover^{son}_{i,t+1}$ is an indicator variable equal to one if the son move away from the county where he grew up and $Promixity_{i,t}$ is the proximity to the nearest train station. The sample includes brothers who are 40-52 years old and their father is observed 30 years earlier. Following equation 5, all regressions include family fixed effects. The instrument is the proximity to DLCP network interacted with $Mover^{son}_{i,t+1}$. Standard errors clustered at the parish level are reported in parentheses. F-stat reports the F-statistic from (Sanderson and Windmeijer, 2016).

Table 6: Spatial Mobility

	(1)	(2)
	$\Pr(\text{Mover}^{son}_{i,c,t+1})$	
Sample	All	Brothers
$\text{Proximity}_{i,c,t}$	0.146	0.161
	(0.030)	(0.036)
F-Stat	135.680	116.947
Observations	980,848	342,715
Avg. dep. var.	0.315	0.293

Note: The coefficients represent standardized $\text{Proximity}_{i,c,t}$, instrumented by the proximity to the DLCP network. The dependent variable is an indicator variable which switches to one if the son moved away from the county where he grew up. All regressions include county and year fixed effects. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its population and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway; the number of servants, household size and whether the father is born outside England and Wales. The sample includes sons (column 1) and brothers (column 2), who are aged 40-52 years old in 1881 or 1911, and their father, observed 30 years earlier. Standard errors clustered at the parish level are reported in parentheses. F-Stat reports the F-statistic from (Sanderson and Windmeijer, 2016).

Table 7: Local Labor Markets Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Urbanization	Industrialization		Skill Level		Social Structure	
	Δlog(pop)	Chimneys	Entrepreneurs	High Skilled	Literacy	Gini	Median
Proximity$_{p,1851}$	0.174	2.789	3.415	0.026	0.174	0.020	-0.157
	(0.064)	(0.457)	(0.374)	(0.004)	(0.020)	(0.004)	(0.510)
Obs.	12,329						
F-Stat	648.715						

Note: Proximity$_{p,1851}$ is the standardized proximity between the parish centroid and the closest rail station in 1851, instrumented by the proximity to the nearest DLCP network. The dependent variable is parish population growth between 1851 and 1881 (column 1), the number of industrial chimneys in a parish (column 2), the number of entrepreneurs per 100m2 (column 3), the Gini of the occupational rank (column 4) and the meadian occupation rank (column 5). All regressions include county and year fixed effects. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its population and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway. F-stat reports the F-statistic from Sanderson and Windmeijer (2016). Standard errors clustered at the parish level are reported in parentheses

A Data Construction

A.1 Data Sources

Census data The I-CeM project, lead by Professor Kevin Sch urer and Professor Eddy Higgs, digitalized and standardized, and coded the England and Wales census of 1851, 1861, 1881, 1891, 1901 and 1911. The I-CeM data collection is available to academic researchers and teachers via the UK Data Service (UKDS) in two forms – in an anonymized version available online `https://icem.data-archive.ac.uk/`, and in a full version (with full name and address) via secure data access arrangements (Schürer and Higgs, 2020).

Occupational ranking HISCAM provides occupational ranks for both national and universal scales. The national scale has been computed using data from Great Britain and is constant for the 1800-1938 period.[27] For the universal scale, however, there is two different candidate scales provided. One that is constant over the same period and another that varies between 1800-1890 and 1890-1938 (Lambert et al., 2013).

HISCO occupation categories The Historical International Standard Classification of Occupations (HISCO) created an occupational classification by customizing the ISCO68 (1968 International Standard Classification of Occupations) to a historical setting (Leeuwen et al., 2002). The occupations are classified by economic sectors and workplace tasks.

Woollard occupation categories The Woollard classification takes into account the class scheme, the status and the division into economic sectors (Woollard, 1998).

HISCLASS occupation categories The HISCLASS classification categorizes occupations into 12 groups based the skill level ranging from unskilled farm workers to higher professional (Van Leeuwen and Maas, 2011). We aggregate these groups into four larger groups: farmers, higher managers, skilled workers, and lower skilled workers. "Farmers" include all agriculture-related activities, "Higher managers" include for instance accountants, solicitors, and clergymen, "Skilled workers" include carpenter, blacksmith, butchers and bricklayers, and "Lower skilled workers" include general laborers, coal miners, or drivers.

Literacy by occupation Using job adverts published in 19th century English periodicals, as well as other contemporaneous descriptions of occupations, Mitch (1992) estimates each occupation group's use of literacy, specifying four categories of jobs: "literacy required"; "literacy likely to be useful"; "possible (or ambiguous) use of literacy"; and "unlikely to use literacy" (Armstrong, 1972).

[27] More information about the computation of the scales can be found at `http://www.camsis.stir.ac.uk/hiscam/`.

Great Britain Address (GB1900) provided by the UK Data Service.

Parish and county boundaries provided by the UK Data Service.

Railways of Great Britain GIS shapefiles of railways lines and stations from 1851 and 1881 from England, Wales and Scotland, digitized by the Cambridge Group for the History of Population and Social Structure. This was digitized from Michael Cobb's definitive atlas The Railways of Great Britain. For more details see the project on Transport, urbanization and economic development in England and Wales c.1670-1911 at `http://www.campop.geog.cam.ac.uk/research/projects/transport/`.

Urban Population data for England and Wales, 1801-1911 from the UK Data Archive Study 7154 (Bennett, 2012). This data collection uses Census returns to construct a consistent time series of population for urban centres in England and Wales 1801-1911.

SRTM Slope DEM for Great Britain. The slope map was created from level 1 SRTM NASA data which was cleaned and had holes patched using a basic nearest neighbour approach and a digital terrain model. This dataset was first accessioned in the EDINA ShareGeo Open repository on 2010-06-30 and migrated to Edinburgh DataShare on 2017-02-20 (Pope, 2017).

Database of historic ports and coastal sailing routes in England and Wales (Alvarez-Palau and Dunn, 2019)

Roman Roads (version 2008) GIS shapefile reflects DARMC's information about the Roman road network identified in the Barrington Atlas (McCormick et al., 2013).

Navigable waterways GIS shapefile on navigable waterways (Satchell and Shaw-Taylor, 2018).

Chimney data (Heblich et al., 2021)

The British Business Census of Entrepreneurs (BBCE) The BBCE is the primary source for large scale information on the business population of entrepreneurs in Victorian and Edwardian Britain. It identifies every self-employed person listed in the censuses for England and Wales 1851-1911, and Scotland 1851-1901, and their employment status as employer, proprietor with no employees, or company director. Available from the UK Data Archive Serve (SN 8600) `https://www.bbce.uk`

A.2 Linking Generations Across Censuses

We create a data-set containing three generations covering the second industrial revolution in Great Britain using the 1851, 1881 and 1911 censuses. Departing from the I-CeM census data, our first step is to link individuals across censuses, so we can later measure fathers' occupations when the son was a child. With this aim, we follow Abramitzky et al. (2019). We use three key variables that do not change over time: year of birth, place of birth and name. The I-CeM provides three variables for the place of birth: county of birth, standardized parish of birth, and parish of birth.

We first standardise names. We then identify potential matches between censuses if (i) the distance between names is smaller than 0.1 bazed on Jaro-Winkler Jaro (1989); Winkler (1999), (ii) the year of births are to be within a ± 2-year window, (iii) they have a perfect match on the place of birth. A match is kept if it is unique and the second best match is far enough in term of year of birth (i.e. if the difference in age between both potential matches is greater than 0). We then apply the data set uniqueness requirement. Specifically, there should be no other person with similar names within his own census. We repeat this for each variable relating to place of birth. The table below presents the number of cases we have.

At the end of the linkage process we have three datasets, one matched based on county of birth, one based on standardized parish and one based on un-standardized parish. We combine these datasets as follows. On a first step we append matches based on standardized and un-standardized parish of birth and find unique pairs. As a result of this step some individuals may not have unique match candidates. Thus we re-apply the selection criteria used above resulting into a dataset containing a unique match per individual. To these data, we add linked observations based on county of birth as long as none of the individuals in the pair is already contained in the parish of birth linked dataset. The resulting dataset contains unique pairs across the three Census years.

A.3 Liking Family Members

Once we have linked individuals across censuses, we link family members. We do this using the within household father identifier provided in the I-CEM data. Thus we are able to link family members even in those cases where we haven't been able to link any individual within the family across censuses. Nonetheless, our interest is on those families where at least a father or a son has been linked across censuses. This is because we want to measure the occupation of the father when the son was young. For this, we need to either have linked the father, the son or both across censuses. In cases where we have only linked the father it must be the case that the son is still living with him. For example, in 1911 Albert Smith, 40, was living with his father John Smith, 60. We were able to link John Smith in 1881 but we have no linkage for Albert Smith. Nonetheless, we do not need this last linkage. As long as we have matched John Smith we are able to observe both his occupation when his son

Table A.1: Linkage Statistics

	County	Std.Parish	Parish
		1851-1881	
Step 1	4,183,316	2,167,519	1,858,074
Step 2	831,566	1,432,347	1,212,960
Step 3	642,315	215,807	171,887
Step 4	1,212,917	1,577,299	1,334,385
Linkage Rate (%)	15	19	16
Linkage Rate Combined (%)		25	
		1881-1911	
Step 1	7,009,691	3,970,122	2,787,243
Step 2	1,539,493	2,629,796	1,916,198
Step 3	1,101,169	430,469	269,908
Step 4	2,151,004	2,909,791	2,098,569
Linkage Rate (%)	17	23	17
Linkage Rate Combined (%)		29	

Note: Step 1 is the number of unique individuals with at least one potential match, Step 2 is the number of unique individuals with unique matches, Step 3 is the number of unique individuals with unique matches after dropping second best match with sufficient age difference, and Step 4 is the number of unique individuals after doing the within cleaning and merging matches from step 2 and step 3. The linkage rate for 1851-1881 (1881-1911) is based on the entire population within the county or parish in 1881 (1911).

was 10 and the occupation of the son 30 years later. Another case, would be that of, for example, Oliver Stone and his father, Harry Stone. We observed both in the 1881 census when Oliver was 12 and the father 35. However, 30 years later, in the 1911 census, we are only able to link Oliver. This case is, again, valid for our analysis as it allows us to observe the occupation of the father when the son was young and the occupation of the son when the son is well into his working life. Obviously any case where we have linked both the father and the son is useful for our analysis. However, any other case outside these three scenarios is not of use for us and we disregard them.

From this set of linked father and sons we keep only those pairs where the son is between

Table A.2: Linkage Statistics for 40-52 years old men

	1851-1881	1881-1911
Nb.Individuals	668,091	1,247,770
Linkage Rate (%)	43	50
Avg.Age Distance	0.56	0.43
Avg.Surname JW-Distance	0.01	0.01
Avg.Name JW-Distance	0.00	0.00

Note: The linkage rate for the 1851-1881 (1881-1911) is based on the population of men aged 40-52 in 1881 (1911).

Table A.3: Comparison with other studies using linked data

Article	Source	Match rate	Number linked
Costas Fernandez et al. (2021)	1881 England and Wales Census to 1911 England and Wales Census (Full, Men 40-52)	50%	1,247,770
Costas Fernandez et al. (2021)	1851 England and Wales Census to 1881 England and Wales Census (Full, Men 40-52)	43%	668,901
Guerra and Mohnen (2020)	1851 London (Full census) to 1881 London (Full, Men 43-49)	33%	263,264
Milner (2019)	1861 England and Wales Census (Full, Men 5-25) to 1881 England and Wales Census (Full, Men 25-45)	37.1%	1,522,047
Milner (2019)	1881 England and Wales Census (Full, Men 5-25) to 1901 England Wales Census (Full, Men 25-45)	42.2%	2,357,948
Long (2005)	1851 England and Wales Census (2% Sample, Men) to 1881 England and Wales Census (Full, Men)	15.2%	28,474
Long and Ferrie (2013)	1881 England and Wales Census (2% Sample, Men 0-25) to 1881 England and Wales Census (Full, Men)	20.3%	14,191
Long and Ferrie (2018)	1881 England and Wales Census (Sons of Men Linked in Long (2005)) to 1911 England and Wales Census (Full, Men)	32.9%	6,672
Feigenbaum (2015)	1915 Iowa Census (Golden & Katz (2000, 2008) Sample, Men 3-17) to 1940 US Census (Full, Men)	57.4%	4,349
Abramitzky et al. (2012)	1865 Norwegian Census (Full, Men 3-15) to 1900 Norwegian Census (Full, Men) or 1900 Roster of Norwegians Immigrants in US (Full, Men)	7.3%	20,446
Abramitzky et al. (2014)	1900 US Census (Subsample of white native & European born men 18-35) to 1910 US Census (Full, Men) and 1920 US Census (Full, Men)	Native Born: 16.5% Immigrant: 8.2%	1,650 20,218
Baker et al. (2018)	1940 US Census (Full, Men born in South 23-58) to 1900, 1910, or 1920 US Census (in each case Full, Men 3-18)	White: 27.5% Black: 18.6%	432,235 170,923

Source: Milner (2020)

40-52 years old. This implies that when the father's occupation was measured, 30 years earlier, the son was 10-22. Moreover, if in any of these father-son pairs has a Jaro-Winkler distance between father and son surname larger than 0.12 we disregard it.

A.4 Geolocating individuals

We geo-locate individuals at two levels: the parish and the address. We geolocate addresses by matching the address provided in the I-CEM data for each individual with the address database put together by the GB1900 team Southall et al. (2017).[28] To improve the quality of the match we split the UK into parishes using the parish identifiers and shape-files provided by I-CEM. In particular, we superimpose parishes on the geo-located addresses and split addresses into disjoint sets according to parish limits. This bounds the error that we can make on geo-locating I-CEM addresses. On a worst case scenario, the distance between the geo-located address and the true address is equal to the maximum distance between two points within the parish and we know that, at least, we are placing the address in the correct parish. After dividing addresses into disjoint subsets by parishes, we make sure that address names are unique within a give parish. If they are not, we have no way to discern between any possible candidate and, therefore, we disregard all non-unique within parish addresses. However, in deciding that an address is unique we introduce some slack. Thus we consider that two seemingly different addresses with the same name are the same if they are no more than 2.5KM away. Then we match address names in the I-CEM data with the geo-located addresses by taking the match with the smallest Jaro-Winkler distance.[29]

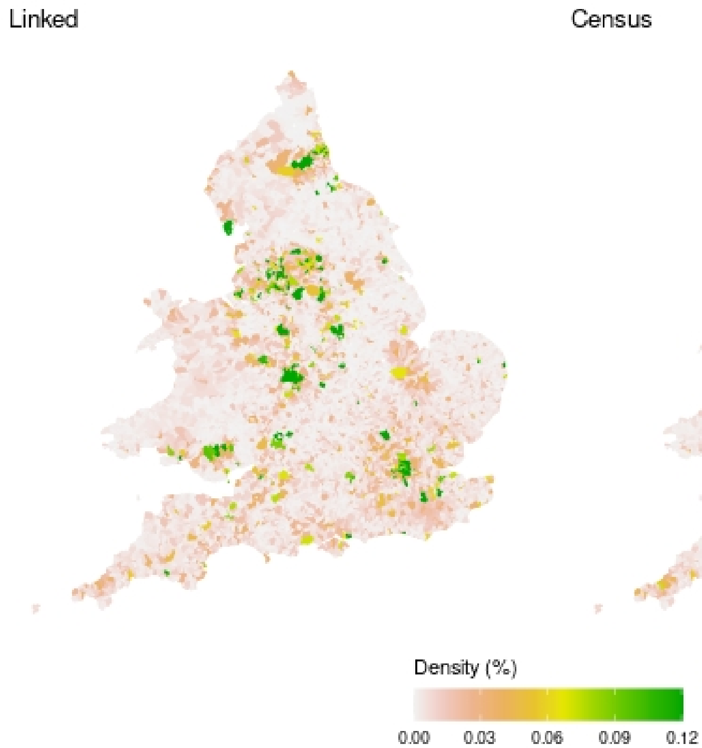
Whenever we use information at the parish level for 1911 we need to standardize the parish definition. This is because the I-CEM data provides a parish division of the UK that is homogeneous for the 1851 and 1881 censuses. However, in the 1911 this division changes. For example, Central London in the 1911 parish division gets divided into five large parishes. We convert the old 1851-1881 parish division into the 1911 division. In most cases, there is a one-to-one mapping (i.e. the 1851-1881 parish is fully contained in a single 1911 parish). Where there is a one-to-many mapping (i.e. the 1851-1881 parish spans multiple 1911 parishes), we split the 1851-1881 parishes by the number of 1911 parishes it spans. To each of these splits we give a weight proportional to share of the original 1851-1881 parish area contained in the split. This was achieved with the GIS files with consistent geographic boundaries (1851-1891 and 1901-1911) provided by Dr. Max Satchell and Dr. Corinne Roughley, both at the University of Cambridge (see `http://www.essex.ac.uk/history/research/icem/documentation.html`.)

[28]The GB1900 final raw gazetteer data dump can be accessed from `http://www.visionofbritain.org.uk/`

[29]A further refinement that one could apply is to also condition on a minimum distance between first and second best match candidate.

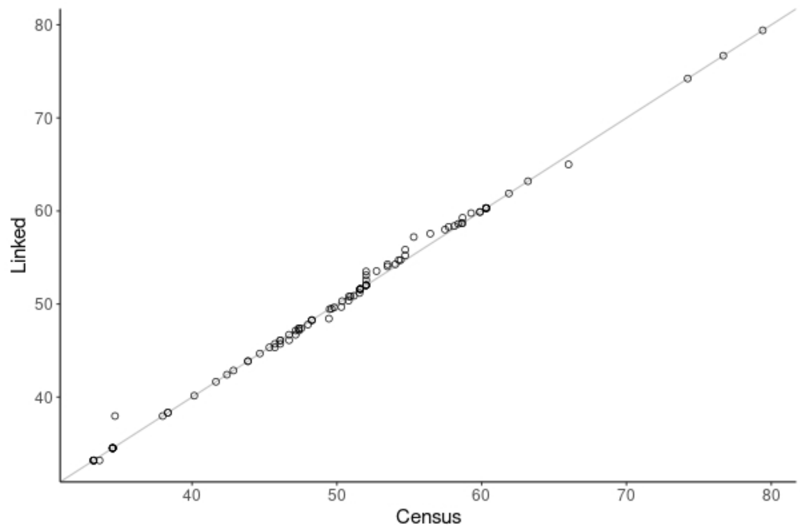
A.5 Descriptives of linked sample

Figure A.1: Size of the sample and population



Note: The figure to the left displays the sample sizes in our main dataset (i.e. linked males that are 40-52 year old in 1881 or 1911) at the parish where they lived 30 years earlier. The figure to the right displays the parish populations of males aged 10-22 pooling data from 1851-1881. Sizes are represented as percentage of the total population.

Figure A.2: Occupational ranking



Note: The dots represent the 1 to 99 percentiles in our estimation sample against the same quantiles in the census for males aged 40-52. Both distributions are constructed by pooling the 1881 and 1911 censuses.

Table A.4: Role of railroad network access on linked sample

Dep. var.: Share of linked individuals among the parish population aged 40-52		
	DLCP network	Nearest train station
	(1)	(2)
Proximity$_{p,c,t}$	0.013	0.002
	(0.007)	(0.004)
Obs	22,041	

Notes: Each coefficient represents the coefficient of the standardised Proximity$_{p,c,t}$ between the parish centroid and the DLCP network (column 1) and between the parish centroid and the nearest train station (column 2). The dependent variable is the share of linked individuals among the parish population aged 10-22 (i.e. sons). All regressions include county and census year fixed effects. Additional controls include the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance "historical importance of town", the distance to the closest Roman road and port "historical travel routes". Standard errors clustered at the parish level are reported in parentheses.

B Additional Descriptives

Table B.1: Mobility Matrix

	Father				
	Managers	Skilled Workers	Lower Skilled	Farmers	Total
Son	Connected				
Managers	**0.453**	0.191	0.175	0.146	176,430
	{ 58,206}	{ 38,315}	{ 58,605}	{ 21,304}	
Skilled Workers	0.177	**0.399**	0.183	0.123	182,242
	{ 22,744}	{ 80,190}	{ 61,304}	{ 18,004}	
Lower Skilled	0.317	0.367	**0.589**	0.348	362,252
	{ 40,797}	{ 73,576}	{197,171}	{ 50,708}	
Farmers	0.053	0.043	0.053	**0.383**	89,233
	{ 6,865}	{ 8,667}	{ 17,841}	{ 55,860}	
Total	128,612	200,748	334,921	145,876	810,157
	Non-connected				
Managers	**0.395**	0.153	0.145	0.121	26,796
	{ 5,992}	{ 5,003}	{ 5,544}	{10,257}	
Skilled Workers	0.184	**0.464**	0.173	0.102	33,230
	{ 2,794}	{15,222}	{ 6,634}	{ 8,580}	
Lower Skilled	0.262	0.273	**0.516**	0.276	56,011
	{ 3,982}	{ 8,947}	{19,746}	{23,336}	
Farmers	0.159	0.110	0.166	**0.501**	54,654
	{ 2,405}	{ 3,604}	{ 6,344}	{42,301}	
Total	15,173	32,776	38,268	84,474	170,691

Note: The entries (in brackets) represent the share (the number) of sons working in a row occupation among sons whose fathers was working in a column occupation. Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. Sons are "connected" if they grew up within 5km of a train station and are "non-connected" if they grew up further than 5km from a train station. The total mobility is 52% for (31% upward mobility and 21% downward mobility) connected sons and 51% for (35% upward mobility and 16% downward mobility) non-connected sons. Occupation classification is based on Woollard.

Table B.2: Descriptives for Brother Sample

	Mean	St.Dev.	Min.	Median	Max.
			A.BROTHERS		
Age	44.91	3.46	40.00	45.00	52.00
Foreign-born	0.02	0.14	0	0	1.00
Urban resident	0.20	0.40	0	0	1.00
Literate	0.37	0.48	0	0	1.00
Occ. rank	49.95	10.02	28.28	50.38	99.00
Occ. cat^{son} $\neq$ Occ. cat^{father}	0.79	0.41	0	1.00	1.00
— Occ. rank^{son} - Occ. rank^{father} —	7.83	8.29	0	5.59	65.80
Upward mobility	0.17	0.38	0	0	1.00
Downward mobility	0.15	0.36	0	0	1.00
County mover	0.29	0.45	0	0	1.00
Dist. moved | county mover	96.66	93.72	0.04	68.59	622.98
Dist. to nearest train station (in km)	3.12	4.92	0.01	1.52	79.96
			B.FATHERS		
Age	47.35	6.76	20.00	47.00	65.00
Foreign-born	0.04	0.20	0	0	1.00
Urban resident	0.15	0.35	0	0	1.00
Household size	7.59	2.04	0	8.00	19.00
Number of sons	5.48	1.97	0	5.00	17.00
Number of servants	0.15	0.59	0	0	11.00
Literate	0.33	0.47	0	0	1.00
Occ. rank	49.32	9.02	28.28	50.95	99.00

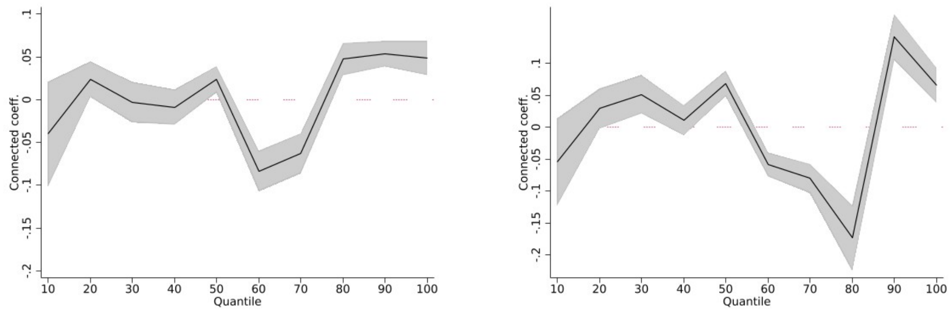
Note: The sample consists of 342,715 sons from 331,932 households. Sons are 10-22 years old when their father's occupation is measured in 1851 or 1881, and 40-52 years old when their own occupation is measured in 1881 or 1911. The table provides descriptives for the sons as adult (panel A) and fathers (panel B).

Table B.3: Change in the share of occupations 1851-1911

Occ. rank		% in 1911	% in 1851
	Top 5 Decreasing		
62110	Farm workers, specialisation unknown	3.76	18.45
61110	General farmers and farmers nfs	1.96	4.47
80100	Boot and shoe makers and repairers	1.39	3.56
75400	Weavers	0.86	2.39
79120	Tailors and tailoresses	0.75	1.90
	Top 5 Increasing		
98550	Delivery men and drivers of goods	2.30	1.34
84130	Machine makers, builders and fitters	1.62	0.20
41010	Dealer, merchant etc. (Wholesale and retail trade)	6.72	4.79
39310	Office clerks, specialisation unknown	3.36	0.78
71120	Miners	7.48	4.26

C Additional Results

Figure C.1: Effect of railroad connection on occ. ranking by percentile



(a) Sons (b) Fathers

Note: Each dot represent the coefficient of the standardized Proximity$_{i,c,t}$ to the nearest train station, instrumented by the proximity to the DLCP network. The shaded region reflects the 95% confidence interval. In figure a (b), the dependent variable is an indicator variable which switches to one if sons (fathers) work in a specific quantile of the HISCAM occupation rank. Observations include sons who are 40-52 years old (figure a) and their fathers (figure b). All regressions include fixed effects for census year and county. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road and port, the number of servants, household size and whether the father is born outside England and Wales.

Table C.1: White vs blue collar occupations

	(1)	(2)
	Father in white collar occ.	Father in blue collar occ.
Occ. cat^{son} $\neq$ Occ. cat^{father}	0.059	0.119
	(0.024)	(0.025)
$\lvert$Occ. rank^{son} - Occ. rank$^{father}\rvert$	-0.574	1.596
	(0.784)	(0.289)
Upward Mobility	0.033	0.069
	(0.018)	(0.015)
Downward Mobility	-0.026	0.015
	(0.031)	(0.009)
F-Stat	69.126	150.878
Obs.	170,305	810,543

Note: Each coefficient represents the coefficient of the standardised Proximity$_{i,c,t}$ to the nearest train station, instrumented by the proximity to the DLCP network. The dependent variable is an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the Occ. rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. Sample is further restricted based on the type of occupation held by the father: white collar (column 1: Occ.cat 0 to 5) and blue collar (column 2: Occ. cat 6 to 9). All regressions include fixed effects for census year and childhood county$_t$. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway; the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish level are reported in parentheses. F-stat reports the F-statistic from Sanderson and Windmeijer (2016).

Table C.2: Intergenerational Mobility Pattern for Brothers

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
Occ. cat^{son} $\neq$ Occ. cat^{father}	0.056	0.026	0.025	0.158	0.126	0.120
	(0.004)	(0.004)	(0.004)	(0.016)	(0.031)	(0.031)
$\lvert$Occ. rank^{son} - Occ. rank$^{father}\rvert$	1.013	0.651	0.639	1.931	1.317	1.239
	(0.057)	(0.057)	(0.056)	(0.210)	(0.413)	(0.409)
Upward Mobility	0.041	0.030	0.029	0.062	0.037	0.034
	(0.003)	(0.003)	(0.003)	(0.010)	(0.019)	(0.019)
Downward Mobility	0.013	0.007	0.006	0.038	0.036	0.035
	(0.002)	(0.002)	(0.002)	(0.008)	(0.016)	(0.016)
Obs.	342,715					
County	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Sub-sample of siblings. Each cell represents the coefficient of the standardised Proximity$_{i,c,t}$ to the nearest train station, instrumented by the proximity to the DLCP railroad network. The dependent variable is an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the Occ. rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include brothers who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county$_t$. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2 and 3); household characteristics including the number of servants, household size and whether the father is born outside England and Wales (column 3). Standard errors clustered at father parish in parenthesis. F-stat reports the F-statistic from Sanderson and Windmeijer (2016)

Table C.3: Growing/declining occupations

	(1)	(2)	(3)
		Occ. of father	
Occ. son	Growing	Declining	All
Growing	0.071	0.307	0.610
	(0.025)	(0.026)	(0.030)
Declining	-0.089	-0.188	-0.588
	(0.024)	(0.022)	(0.027)
Obs.		980,848	

Note: Growing/declining is an indicator variable is an individual works in a Occ. cat within the top/bottom 25% of the growth industry (see Table B.3 for examples). The growth of industry is based on the difference in the share of individuals in a Occ. cat between 1851 and 1911. All regressions include fixed effects for census year and childhood county$_t$. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway; the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish level are reported in parentheses.

Table C.4: The effect of rail connection by HISCLASS occupations

	(1)	(2)	(3)	(4)	(5)
			Father		
Son	Manager	Skilled Workers	Lower Skilled	Farmers	All
Manager	0.058	0.036	0.032	0.035	0.037
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Skilled Workers	0.003	-0.048	0.026	0.028	0.003
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Lower Skilled	0.004	0.045	0.015	0.050	0.033
	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)
Farmers	-0.066	-0.034	-0.072	-0.113	-0.073
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Obs.			980,848		

Note: Each entry represents the coefficient of the standardised Proximity$_{i,c,t}$ to the nearest train station instrumented by the proximity to the DLCP network. The dependent variable is an indicator equal to one if the son works in a specific HISCLASS occupation (rows). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier (column 5). Additional sample restriction are that the fathers work as "manager" (column 1), "skilled worker" (column 2), "lower skilled worker" (column 3), and "farmer" (column 4). All regression include census year and childhood county$_t$ fixed effects. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway; the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at father parish in parenthesis.

Table C.5: Skill level based on occupation

	(1)	(2)	(3)	(4)
	Father			
Son	Manual	Non-Skilled	Non-Managerial	All
Non-Manual	0.049	0.068	0.058	0.073
	(0.012)	(0.015)	(0.013)	(0.015)
Skilled	0.019	0.020	0.022	0.024
	(0.005)	(0.005)	(0.005)	(0.005)
Managerial	0.015	0.029	0.019	0.030
	(0.010)	(0.012)	(0.010)	(0.012)
Obs.	980,848			

Note: Each entry represents the coefficient of the standardized proximity, instrumented by the proximity to the DLCP network in a control function approach. The dependent variable is the whether the son is literate (row 1) and whether the son is skilled (row 2). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. Average effects reported in column 4. We also report estimates by father skill: manual (column 1, 837,063 obs.), non-skilled (column 2, 949,469 obs.), an non-managerial (column 3, 861,000 obs.). All regressions include county and year fixed effects. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway; the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at parish level in parenthesis.

Table C.6: Decomposition

	(1)	(2)	(3)	(4)
	Occ. cat^{son} $\neq$ Occ. cat^{father}	—Occ. rank^{son} - Occ. rank^{father}—	Upward Mobility	Downward Mobility
Promixity	0.120	1.239	0.034	0.035
	(0.031)	(0.408)	(0.019)	(0.016)
Local opportunities	0.149	1.493	0.035	0.043
	(0.032)	(0.416)	(0.019)	(0.017)
Ease of spatial mobility	0.007	0.127	0.005	0.000
	(0.002)	(0.034)	(0.001)	(0.001)
Returns to spatial mobility	-0.036	-0.381	-0.006	-0.009
	(0.004)	(0.068)	(0.003)	(0.003)
Obs.		342,715		

Note: The sample consists of brothers who are 40-52 years old and their father is observed 30 years earlier.
"Proximity" $= \Delta \Pr(IM|Train) =$ Total ($\hat{\alpha}$) (see Table C.2)
"Local opportunities" $= \Delta \Pr(IM|Stayer, Train) = \hat{\alpha}$ - $\hat{\tau}\hat{\phi}$ - $\hat{\lambda}\frac{\sum_t \sum_{i=N}^{N_t} Mover_{i,f,t}}{\sum_t N_t}$
"Ease of spatial mobility" $= [\Pr(IM|Mover, Train) - \Pr(IM|Stayer, Train)] \times \Delta \Pr(Mover|Train) = \hat{\tau}\hat{\phi}$
"Returns to spatial mobility" $= [\Delta \Pr(IM|Mover, Train) - \Delta \Pr(IM|Stayer, Train)] \times \Pr(Mover|Train) = \hat{\lambda}\frac{\sum_t \sum_{i=N}^{N_t} Mover_{i,f,t}}{\sum_t N_t}$
Notation as introduced in equations (1), (5) and (6). Estimates needed for the decomposition are obtained by estimating a system of equations through GMM. Standard errors clustered at parish level in parenthesis.

D Robustness Checks

Alternative definition of connectedness In our baseline specification, we define connectedness based on the distance to the nearest train station. We explore alternative measures of connectedness defined as (1) an indicator variable equal to one if the son grew up within 5, 10 and 15km of a train station, (2) an indicator variable equal to one if the son grew up with a train station within his parish borders, and (3) distance to the railroad network. Figure D.1 shows that our baseline results are conservative.

Alternative measure of intergenerational mobility We also examine how sensitive our results is to the HISCAM occupation ranking and alternative measures of upward and downward mobility in Figure D.2. We first use the HISCAM occupation ranking that takes into account changes in the ranking of occupations over time (e.g. being a farmer in 1851 may not reflect the same prestige as being a farmer in 1881 (Xie and Killewald, 2013)). As a second alternative occupation ranking, we use 0.5, 1.5 and 2 standard deviation instead of the 1 standard deviation in the baseline for the definitions of upward and downward mobility. In all cases, our results remain robust to these alternative measures of intergenerational mobility. Results are not statistically different from other measure of mobility.

Rail related occupations Railroad came with specific occupations such as train conductor or controller. Better connected areas would mechanically employ more residents in such positions. We therefore remove any occupations related to the railroad in Table D.1. We see that the our results are robust.

Alternative specification Figure D.3 shows the results once we add higher polynomials to the control variables and parish fixed effects. There are 10,419 parishes and consequently the parish fixed effect controls for very local characteristics such as public good provisions, the initial wealth and local industries. Unsurprisingly, when including parish fixed effect, the effect of proximity to the railroad network on occupational ranking becomes smaller in magnitude except in the case of changes in occupational categories between fathers and sons where the effects are similar in magnitude.

Alternative instruments The instrument used in our baseline specification exploits both the spatial and temporal variation in the expansion of the railroad network. Figure D.4 presents alternative instruments. The first creates the least cost path network where the cost is solely based on distance and not the slope of the terrain. The second creates a static least cost path that does not distinguish between early and late lines.

Parish level location There may be measurement error in the location of individual within a parish given the string matching between street address reported in the census and the geocoded street names. This would affect the measure of connected in our baseline specification defined as the distance between the residence and the nearest railroad station. As a robustness check, we use the parish centroid as the location of individuals. We then measure connectedness based on the distance between the parish centroid and the location of the nearest railroad station. In Table D.2 we see that our results are robust to potential measurement error.

Linking procedure A primary concern in creating intergenerational mobility is the false positives (i.e. linking children to the wrong parents). Moreover, the linked sample may suffer from selection problems. In particular, it is likely that families that stay in England and Wales more stable are overrepresented. Furthermore, people, belonging to the middle class and with higher education, are more likely to be able to accurately answer the census questions. If individuals in connected areas are more likely to move and/or acquire higher level of education, our mobility rates may be biased. Given that we no do observe the outcomes and connectedness to the railroad network of non-linked individuals, we proxy the probability of linkage using the proportion of linked individuals within county-of-birth, census year and first name frequency. We do not use surname frequency as this has been shown to be correlated with wealth. In Table D.3, we control for the probability of being linked using a polynomial.

Heterogeneity effects by distance Our IV estimates identify a local average treatment effect among the set of compliers. Here, the compliers are individuals residing close to a train station because their location is convenient placed close to the DLCP network but would not have been close otherwise. In the presence of continuous, endogenous, and heterogeneous treatment effects, our linear IV estimate identifies a weighted average of the underlying marginal causal effects across the proximity distribution (Angrist and Imbens, 1995). The weight attached to each value of proximity depends on the proportion of sons who, because of the instrument, experience a change in proximity to the nearest train station. Hence more weight is given to the marginal effects for proximities that are most affected by the instrument (proximity to the DLCP). To understand the relate contribution of each observation to our IV estimate, we compute the causal response weighting function following the decomposition proposed by Løken et al. (2012). To do so, we allow the proximity to the railroad to take discrete jumps of Δ meters. Call $\mathrm{DProx}_{d,i,c,t-1} = 1\left\{\mathrm{Proximity}_{i,c,t-1} \geq d \times \Delta\right\}$ where $d \in$

$\{0, 1, \ldots, \overline{d}\}$ such that $\max \text{Proximity}_{i,c,t-1} \leq \overline{d} \times \Delta$. The unrestricted IV model is

$$f(Rank^{son}_{i,c,t}, Rank^{father}_{i,c,t-1}) = \sum_{d=1}^{\overline{d}} \beta_d \text{DProx}_{d,i,c,t-1} + \gamma_t + \rho_c + \nu_{i,c,t-1}$$

Løken et al. (2012) show that

$$\alpha_1^{IV} = \sum_{d=1}^{\overline{d}} w_d^{IV} \beta_d,$$

where

$$w_d^{IV} = \frac{Cov\left(\text{DProx}_{d,i,c,t-1}, \text{Proximity to DLCP network}_{i,c,t-1}\right)}{Cov\left(\text{Proximity}_{i,c,t-1}, \text{Proximity to DLCP network}_{i,c,t-1}\right)}.$$

In Figure D.5 we report the causal response weighting function w_d^{IV} and the population distribution of proximity to the nearest train station. We see that we have compliers across the entire distribution of proximity. The weights that the IV linear estimation assigns to the marginal effect are highest for individuals residing within 0.5 and 1.5 proximity units (i.e., within 2.7 and 8.1km to a train station). These individuals are the ones whose proximity to the railway are most affected by being along the hypothetical railroad network path. The highest weights do not coincide with the distribution to the proximity in our sample. A large proportion of our sample live less than 5.4km away from a train station. Unsurprisingly, these individuals contribute to our IV but do not contribute the most since they tend to live close to town centres and would have been close to the train station regardless of our instrument.

To understand how the linearity assumption affects our results, we run the following quadratic specification:

$$\begin{aligned} f(Occ^{son}, Occ^{father})_{i,c,t} = \theta_1 Proximity_{i,c,t} + \theta_2 (Proximity_{i,c,t})^2 \\ + \gamma_t + \rho_c + \epsilon_{i,c,t} \end{aligned} \tag{7}$$

We use the square distance to the hypothetical railroad network as an instrument for $(\text{Proximity}^2_{i,c,t})$. Figures D.6 present the predicted marginal effects for our four outcome variables. The closer to the train station, the larger the effects of proximity on intergenerational mobility, which suggests non-linear effects.

Year Table D.4 splits the sample by census year. We see that the intergenerational mobility patterns remain in both subsamples, although the magnitudes are larger in the later period.

Excluding one region at a time We show that the results are robust to excluding one county at a time. Figure D.7 shows that our findings are not confined to a single region.

Urban vs. rural We examine the effect of the railroad network on the intergenerational mobility patterns for sons who grew up in an urban (i.e. those who grew up within 2.5km of a 1801 town) and non-urban areas. In Table D.5, we do not observe large differences between the two groups.

Removing individuals at nodes A potential concern is that our result are mainly driven individuals residing at the nodes of our railroad network. In Table D.6 we remove individuals within 2.5km of 1801 major towns (i.e. the nodes of our network). Our results remains robust thereby alleviating concerns related to urban centres.

Age As several studies have shown (e.g. Grawe (2006)), estimates of intergenerational mobility is highly sensitive to the age at which sons' labor market outcomes are observed, increasing substantially in age. This can be explained by the strong life-cycle pattern in the correlation between current and lifetime earnings. In the baseline sample, fathers are between 20 and 65 years old and their sons are between 10 to 22 years old. Older fathers may be more likely to be established in their profession and provide a financially stable environment for their sons. In Table D.7 we do not see differences in the effects of having access to the railroad network by the age of the father. In the baseline, we measure connectedness during youth when the sons lived with their fathers. Similarly, we look at the age of sons in Table D.8. We restrict the sample of sons by their age to account for the fact that younger sons have not chosen their occupation and can therefore benefit from the new opportunities brought by the railroad network. We see the effects of being better connected as similar no matter the ages of sons. The only difference is for sons aged 17 to 22 for which being better connected has a positive and significant effect on downward mobility.

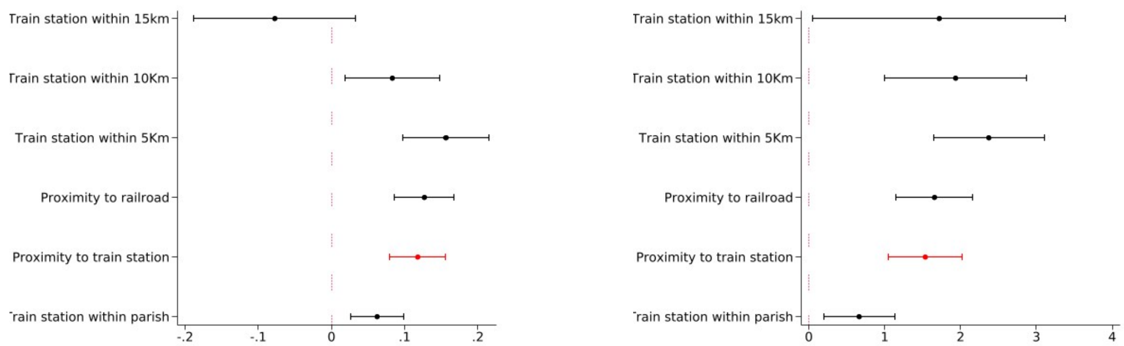
Birth order The birth order of sons may play a role in the intergenerational mobility patterns if, for instance, first-born sons inherit family businesses. Table D.9 shows that the effect of the railroad network affected all sons, regardless of their birth order.

Natives vs. foreigners Recent work by Abramitzky et al. (2012); Abramitzky, Boustan and Eriksson (2014) shows that migration status is an important factor for intergenerational mobility patterns. In Table D.10 we separate the sample of native, first and second generation sons and examine the effect of the access to the railroad network on their intergenerational mobility pattern. We find that our results are mainly driven by natives. We also see that better connected foreigners experienced large upward mobility.

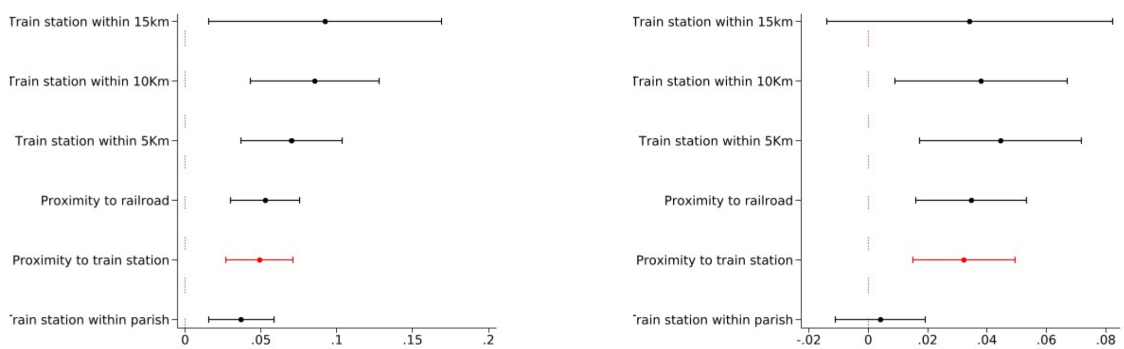
Locals vs. outsiders The estimator would also be biased if people and firms move over time along the same spatial lines as the forecasted placement of the railroad network. For instance, fathers who have high ambition for their family may decide to live in connected parishes. We explore the possibility of self-selection in two ways. In Table 5, we investigate the differences in the effect of access to the railroad network on sons who grew up in a different county as the one they are currently living in (i.e. movers). In Table D.11 we also examine the effects for sons who were born in a different county from the one they grew up in (i.e. outsiders)

Farming activities Railroads had a major impact on farming, as perishable goods such as dairy products could now be moved long distances before they were inedible. In Table D.12 we split the sample between fathers who are in farming activities and all other activities. We see that the general intergenerational mobility patterns are robust to this split.

Figure D.1: Alternative definition of proximity



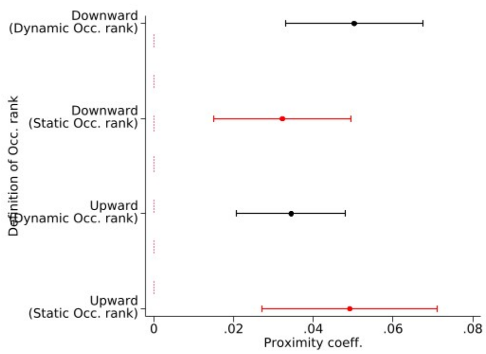
(a) Occ. $\text{cat}^{son} \neq$ Occ. cat^{father} (b) |Occ. rank^{son} - Occ. rank^{father}|



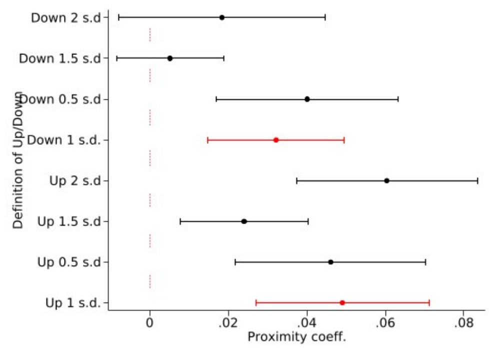
(c) Upward mobility (d) Downward mobilty

Notes: Each dot represents the coefficient of the standardized $\text{Proximity}_{i,c,t}$, instrumented by the proximity to the DLCP network. Proximity is defined as the distance to the nearest train station (red dot), indicator if the parish has a train station (first black dot), indicator if the train station is within 15/10/5 km, or the distance to the nearest railroad (last black dot). The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (Figure a), the absolute value of the difference in the HISCAM occupational rank between sons and fathers (Figure b), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (Figure c / Figure d). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include census year and county fixed effects and controls for the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road and port, the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. The lines represents the 95% confidence interval. *p<0.1; **p<0.05; ***p<0.01

Figure D.2: Alternative definition of occupation ranking



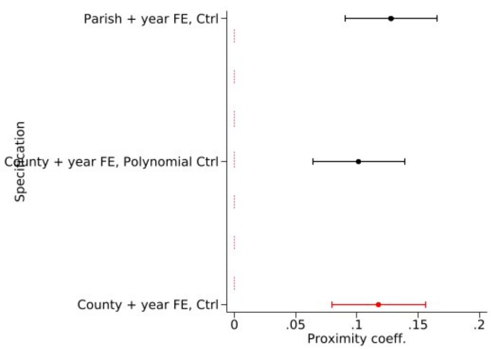
(a) Definition of HISCAM



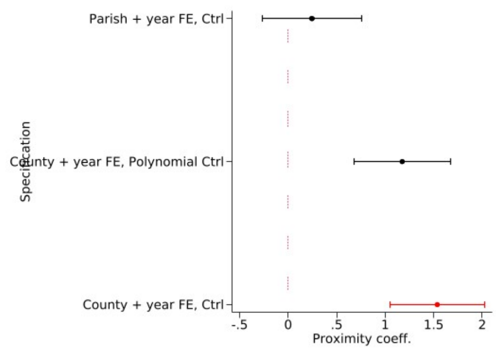
(b) Definition of Up/Down

Notes: Each dot represents the coefficient of the standardized $\text{Proximity}_{i,c,t}$, instrumented by the proximity to the DLCP network. In Figure a, the dependent variable is the absolute value of the difference between father and son in the HISCAM occupational rank (red dot) or the dynamic HISCAM (black dot). In Figure b, the dependent variable is an an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than 0.5, 1, 1.5 or 2 standard deviation. Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include census year and county fixed effects, and controls for the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road and port, the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. The lines represents the 95% confidence interval. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

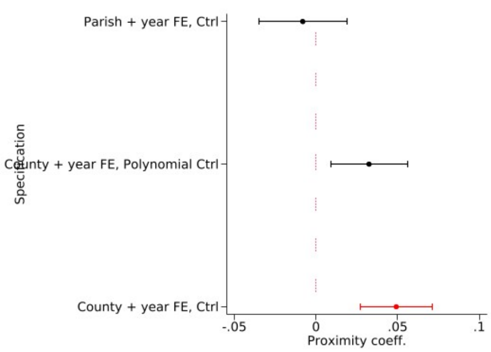
Figure D.3: Alternative specification



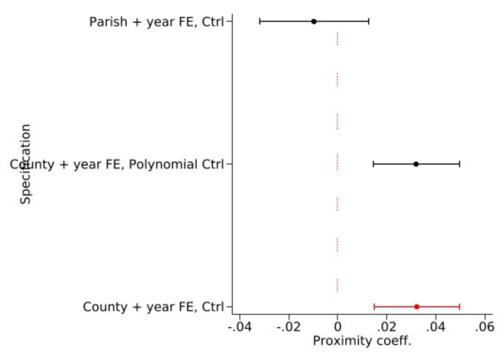
(a) Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}



(b) |Occ. rank^{son} - Occ. rank^{father}|



(c) Upward mobility



(d) Downward mobilty

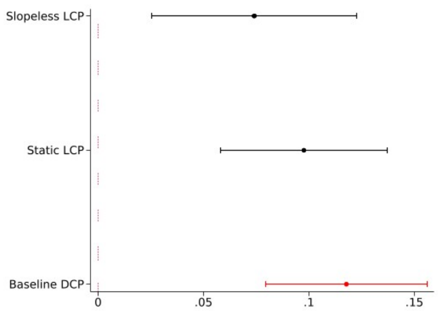
Notes: Each dot represents the coefficient of the standardized $\text{Proximity}_{i,c,t}$, instrumented by the proximity to the DLCP network. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (Figure a), the absolute value of the difference in the HISCAM occupational rank between sons and fathers (Figure b), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (Figure c / Figure d). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. The baseline regression (red dot) includes fixed effects for census year and county, controls for the historical importance of town and historical travels routes and controls for household characteristics. The first black dot also includes parish fixed effects and the second black dot includes second order polynomials for the control variables. Standard errors clustered at the parish level are reported in parentheses. The lines represents the 95% confidence interval. *p<0.1; **p<0.05; ***p<0.01.

Table D.1: Main results without rail related occupations

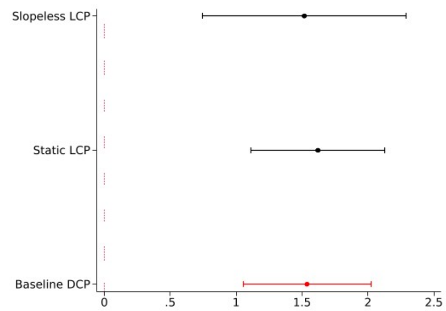
	(1)	(2)	(3)
Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}	0.155	0.119	0.112
	(0.013)	(0.024)	(0.024)
$\vert$Occ. rank^{son} - Occ. $\text{rank}^{father}\vert$	2.012	1.392	1.282
	(0.166)	(0.297)	(0.292)
Upward Mobility	0.066	0.046	0.042
	(0.007)	(0.013)	(0.013)
Downward Mobility	0.041	0.031	0.030
	(0.005)	(0.011)	(0.011)
F-Stat	417.526	134.950	134.937
Obs.		929,638	
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel routes	No	Yes	Yes
Household characteristics	No	No	Yes

Note: The dependent variable is indicated in the row label. The independent variable is the standardised negative distance between the childhood residence and the nearest railroad station instrumented with proximity to the DLCP. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2 and 3); household characteristics including the number of servants, household size and whether the father is born outside England and Wales (column 3). Standard errors clustered at parish level in parenthesis. F-stat reports the F-statistic from Sanderson and Windmeijer (2016).

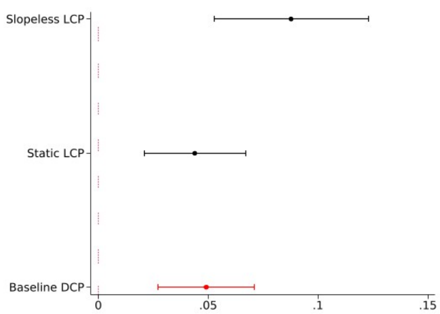
Figure D.4: Alternative instrument



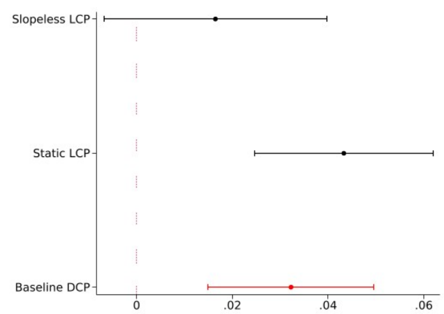
(a) Occ. cat^{son} $\neq$ Occ. cat^{father}



(b) |Occ. rank^{son} - Occ. rank^{father}|



(c) Upward Mobility



(d) Downward mobility

Table D.2: Measurement error in geolocation

	(1)	(2)	(3)	(4)	(5)	(6)
	Distance from Address			Distance from Parish Centroid		
Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}	0.151	0.118	0.111	0.153	0.113	0.107
	(0.013)	(0.023)	(0.023)	(0.015)	(0.024)	(0.024)
$\mid$ Occ. $\text{rank}^{son}-$ Occ. $\text{rank}^{father}\mid$	2.034	1.537	1.429	2.054	1.433	1.336
	(0.164)	(0.295)	(0.289)	(0.171)	(0.264)	(0.260)
Upward Mobility	0.068	0.049	0.045	0.067	0.044	0.040
	(0.007)	(0.013)	(0.013)	(0.007)	(0.012)	(0.012)
Downward Mobility	0.038	0.032	0.031	0.039	0.031	0.029
	(0.005)	(0.011)	(0.011)	(0.006)	(0.010)	(0.010)
F-Stat	414.450	135.729	135.680	366.859	153.789	153.787
Obs.	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

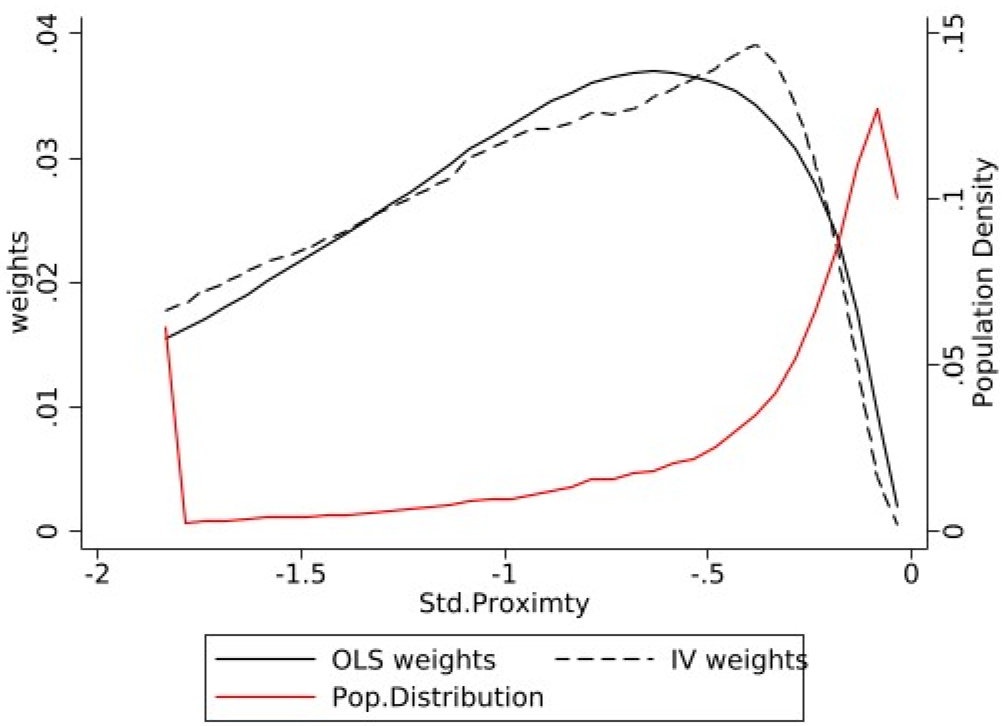
Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$, instrumented by the proximity to the DLCP network. Individuals are geolocated based on their address (columns 1 to 3) or at the parish centroid (columns 4 to 6). The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the Occ. rank rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5 and 6); household characteristics including the number of servants, household size and whether the father is born outside England and Wales (columns 3 and 6). Standard errors clustered at parish level in parenthesis. F-stat reports the F-statistic from Sanderson and Windmeijer (2016).

Table D.3: Controlling for the selection

	(1)	(2)	(3)
Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}	0.151	0.117	0.110
	(0.013)	(0.023)	(0.023)
$\mid$ Occ. rank^{son} - Occ. $\text{rank}^{father} \mid$	2.034	1.532	1.424
	(0.164)	(0.294)	(0.289)
Upward Mobility	0.067	0.049	0.045
	(0.007)	(0.013)	(0.013)
Downward Mobility	0.038	0.032	0.031
	(0.005)	(0.011)	(0.010)
Obs.		980,848	
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel routes	No	Yes	Yes
Household characteristics	No	No	Yes
Cubic Prob.Linkage	Yes	Yes	Yes

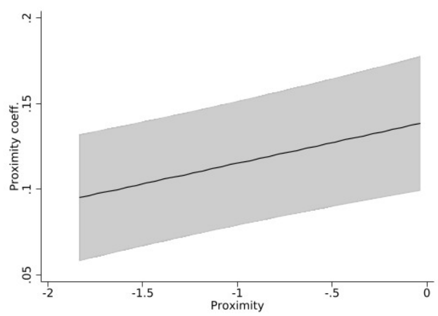
Note: Each cell represents the coefficient of the standardized $\text{Proximity}_{i,c,t}$ to the nearest train station (columns 1 to 4) and instrumented by the proximity to the DLCP railroad network (columns 5 to 8). The dependent variable is an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the occupational rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and county. Additional controls include a cubic polynomial on the probability linkage, dummies for the frequency of the surname, the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5 and 6); household characteristics including the number of servants, household size and whether the father is born outside England and Wales (columns 3 and 6). Standard errors clustered at the parish level are reported in parentheses.

Figure D.5: IV weights

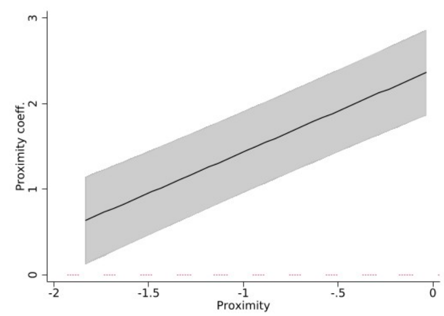


Note: This figure shows the population share (right axis) and the assigned weights in the IV estimates (left axis) over the proximity to the nearest train station. The x-axis represents units (5.4 km each) of proximity to the nearest train station winsorized at the 1%.

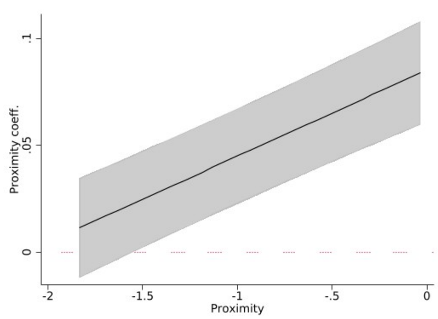
Figure D.6: Predicted marginal effect



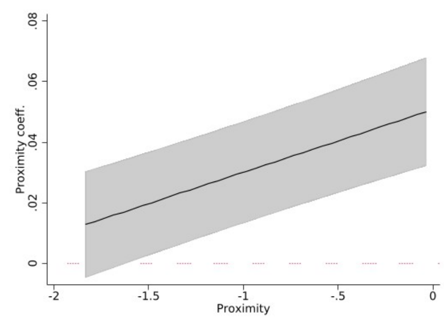
(a) Occ. cat^{son} $\neq$ Occ. cat^{father}



(b) |Occ. rank^{son} - Occ. rank^{father}|



(c) Upward mobility



(d) Downward mobility

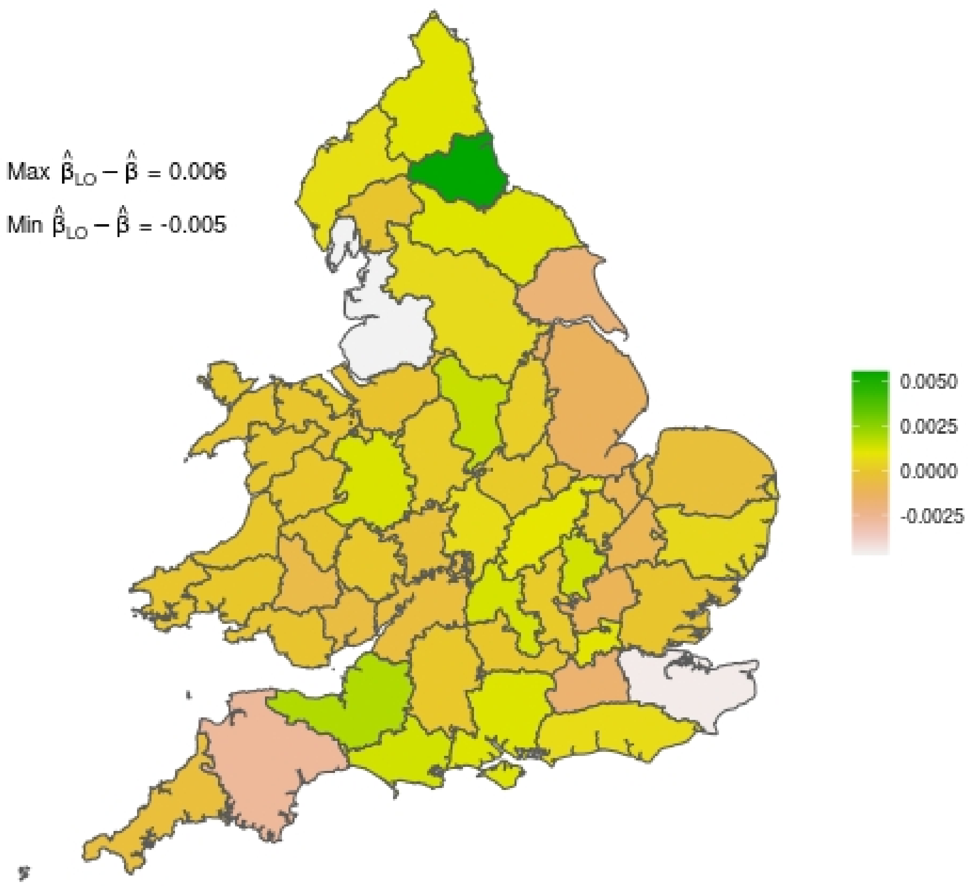
Note: This figure presents the predicted marginal effect of equation 7. The x-axis represents units (5.4 km each) of proximity to the nearest train station winsorized at the 1%.

Table D.4: Subsample by year

	(1)	(2)	(3)	(4)	(5)	(6)
	1851-1881			1881-1911		
Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}	0.123	0.085	0.079	0.170	0.126	0.117
	(0.010)	(0.015)	(0.015)	(0.021)	(0.045)	(0.045)
$\lvert$Occ. $\text{rank}^{son}-$ Occ. $\text{rank}^{father}\rvert$	1.803	1.466	1.373	2.342	1.772	1.626
	(0.144)	(0.250)	(0.244)	(0.262)	(0.539)	(0.532)
Upward Mobility	0.066	0.051	0.048	0.069	0.039	0.034
	(0.006)	(0.011)	(0.011)	(0.012)	(0.026)	(0.026)
Downward Mobility	0.022	0.019	0.017	0.060	0.068	0.066
	(0.005)	(0.009)	(0.009)	(0.009)	(0.022)	(0.022)
F-Stat	397.836	122.981	122.986	221.224	53.016	52.984
Obs.		281,912			698,936	
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$, instrumented by the proximity to the DLCP network. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the Occ. rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include sons who are 40-52 years old in 1881 (columns 1 to 3) and in 1911 (columns 4 to 6) and their father's occupation is measured 30 years earlier. All regressions include fixed effects for childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5 and 6); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3 and 6). Standard errors clustered at parish level in parenthesis. F-stat reports the F-statistic from Sanderson and Windmeijer (2016).

Figure D.7: Excluding one county at a time



Note: We estimate equation 1 excluding one county at a time. The figure plots the coefficient of the standardized Proximity$_{i,c,t}$, instrumented by the proximity to the DLCP network, for each county excluded. Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for county. Additional controls include the historical importance of town, historical travels routes and household characteristics consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road and port, the number of servants, household size and whether the father is born outside England and Wales.

Table D.5: Urban vs. Rural

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. $cat^{son} \neq$ Occ. cat^{father}			$\vert$ Occ. $rank^{son}-$ Occ. $rank^{father} \vert$		
$\text{Proximity}_{i,c,t}$	0.135	0.110	0.104	1.983	1.578	1.492
	(0.012)	(0.022)	(0.022)	(0.163)	(0.289)	(0.285)
Urban $\text{Location}_{i,c,t}$ * $\text{Proximity}_{i,c,t}$	-0.049	-0.031	-0.034	-1.393	-1.186	-1.218
	(0.009)	(0.009)	(0.009)	(0.152)	(0.156)	(0.155)
	(7)	(8)	(9)	(10)	(11)	(12)
	Upward Mobility			Downward Mobility		
$\text{Proximity}_{i,c,t}$	0.069	0.053	0.050	0.033	0.030	0.028
	(0.007)	(0.013)	(0.013)	(0.005)	(0.010)	(0.010)
Urban $\text{Location}_{i,c,t}$ * $\text{Proximity}_{i,c,t}$	-0.056	-0.051	-0.052	-0.017	-0.013	-0.013
	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)
Obs.Total	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$ interacted with whether the father was living in an urban area, where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1-3), the absolute value of the difference in the Occ. rank between sons and fathers (columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (columns 7-9/10-12). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. We have 163,626 urban and 817,222 non-urban observations. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town , its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5, 6, 8, 9, 11 and 12); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3, 6, 9 and 12). Standard errors clustered at parish level in parenthesis.

Table D.6: Excluding individuals at nodes

	(1)	(2)	(3)
Occ. cat^{son} $\neq$ Occ. cat^{father}	0.092	0.050	0.047
	(0.015)	(0.026)	(0.026)
$\mid$ Occ. rank^{son}$-$ Occ. rank^{father}$\mid$	1.657	1.161	1.116
	(0.213)	(0.333)	(0.328)
Upward Mobility	0.058	0.044	0.042
	(0.010)	(0.016)	(0.016)
Downward Mobility	0.027	0.014	0.014
	(0.007)	(0.012)	(0.012)
Obs.		787,599	
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel routes	No	Yes	Yes
Household characteristics	No	No	Yes

Note: Each entry represents the coefficient of the standardised Proximity$_{i,c,t}$, instrumented by the proximity to the DLCP network. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the Occ. rank between sons and fathers (row 2), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (row 3/row 4). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier, without sons who live within 2.5 km of a 1801 major town (at the top 10% of population in 1801). All regressions include fixed effects for census year and childhood county$_t$. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (column 2); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3). Standard errors clustered at parish level in parenthesis.

Table D.7: Age of father

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. $cat^{son} \neq$ Occ. cat^{father}			$\mid$ Occ. $rank^{son} -$ Occ. $rank^{father} \mid$		
$\text{Proximity}_{i,c,t}$	0.136	0.105	0.097	1.883	1.418	1.296
	(0.017)	(0.025)	(0.025)	(0.267)	(0.359)	(0.356)
Father 31-40 * $\text{Proximity}_{i,c,t}$	-0.006	-0.007	-0.006	-0.125	-0.146	-0.134
	(0.012)	(0.011)	(0.011)	(0.211)	(0.211)	(0.211)
Father 41-50 * $\text{Proximity}_{i,c,t}$	0.012	0.010	0.011	0.076	0.049	0.062
	(0.011)	(0.011)	(0.011)	(0.210)	(0.210)	(0.210)
Father 51-65 * $\text{Proximity}_{i,c,t}$	0.035	0.033	0.033	0.410	0.377	0.386
	(0.012)	(0.012)	(0.012)	(0.211)	(0.210)	(0.210)
	(7)	(8)	(9)	(10)	(11)	(12)
	Upward Mobility			Downward Mobility		
$\text{Proximity}_{i,c,t}$	0.062	0.044	0.039	0.041	0.035	0.034
	(0.012)	(0.017)	(0.017)	(0.011)	(0.014)	(0.014)
Father 31-40 * $\text{Proximity}_{i,c,t}$	-0.002	-0.003	-0.002	-0.006	-0.007	-0.007
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Father 41-50 * $\text{Proximity}_{i,c,t}$	0.002	0.001	0.002	-0.003	-0.004	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Father 51-65 * $\text{Proximity}_{i,c,t}$	0.014	0.013	0.013	0.001	0.000	0.000
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Obs.Total	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$ interacted with father age 30 years earlier been 20-30 (baseline, 6,148 obs.), 31-40 (223,862 obs.), 41-50 (456,495 obs.) and 51-65 (294,343 obs.); where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1-3), the absolute value of the difference in the Occ. rank between sons and fathers (columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (columns 7-9/10-12). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town , its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5, 6, 8, 9, 11 and 12); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3, 6, 9 and 12). Standard errors clustered at parish level in parenthesis.

Table D.8: Age of son

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. cat^{son} $\neq$ Occ. cat^{father}			$\mid$ Occ. rank$^{son}-$ Occ. rank$^{father}\mid$		
Proximity$_{i,c,t}$	0.137	0.102	0.095	1.893	1.390	1.279
	(0.012)	(0.022)	(0.022)	(0.166)	(0.291)	(0.287)
Son 12-30 * Proximity$_{i,c,t}$	0.008	0.008	0.008	0.037	0.035	0.039
	(0.002)	(0.002)	(0.002)	(0.046)	(0.046)	(0.046)
Son 14-16 * Proximity$_{i,c,t}$	0.021	0.021	0.022	0.222	0.222	0.224
	(0.002)	(0.002)	(0.002)	(0.043)	(0.042)	(0.042)
Son 17-22 * Proximity$_{i,c,t}$	0.035	0.034	0.035	0.337	0.331	0.335
	(0.002)	(0.002)	(0.002)	(0.042)	(0.042)	(0.042)
	(7)	(8)	(9)	(10)	(11)	(12)
	Upward Mobility			Downward Mobility		
Proximity$_{i,c,t}$	0.065	0.047	0.043	0.037	0.031	0.030
	(0.007)	(0.013)	(0.013)	(0.005)	(0.010)	(0.010)
Son 12-30 * Proximity$_{i,c,t}$	0.001	0.000	0.001	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Son 14-16 * Proximity$_{i,c,t}$	0.005	0.005	0.005	0.000	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Son 17-22 * Proximity$_{i,c,t}$	0.004	0.004	0.004	0.004	0.004	0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs.Total	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised Proximity$_{i,c,t}$ interacted with son age 30 years earlier been 10-11 (baseline, 341,816 obs.), 12-13 (178,270 obs.), 14-16 (223,273 obs.) and 17-22 (237,489 obs.); where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1-3), the absolute value of the difference in the Occ. rank between sons and fathers (columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (columns 7-9/10-12). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county$_t$. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town , its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5, 6, 8, 9, 11 and 12); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3, 6, 9 and 12). Standard errors clustered at parish level in parenthesis.

Table D.9: Birth Order

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. $\text{cat}^{son} \neq$ Occ. cat^{father}			$\mid$ Occ. $\text{rank}^{son} -$ Occ. $\text{rank}^{father} \mid$		
$\text{Proximity}_{i,c,t}$	0.151	0.116	0.109	2.027	1.516	1.405
	(0.013)	(0.022)	(0.022)	(0.166)	(0.292)	(0.288)
Second Born * $\text{Proximity}_{i,c,t}$	0.001	0.001	0.000	0.012	0.008	0.006
	(0.002)	(0.002)	(0.002)	(0.036)	(0.036)	(0.036)
Third Born * $\text{Proximity}_{i,c,t}$	-0.001	-0.002	-0.002	-0.025	-0.033	-0.033
	(0.003)	(0.003)	(0.003)	(0.047)	(0.047)	(0.047)
Fourth and Above Born * $\text{Proximity}_{i,c,t}$	0.003	0.003	0.003	0.014	0.010	0.014
	(0.003)	(0.003)	(0.003)	(0.056)	(0.056)	(0.056)
	(7)	(8)	(9)	(10)	(11)	(12)
	Upward Mobility			Downward Mobility		
$\text{Proximity}_{i,c,t}$	0.068	0.050	0.046	0.037	0.031	0.029
	(0.007)	(0.014)	(0.013)	(0.005)	(0.010)	(0.010)
Second Born * $\text{Proximity}_{i,c,t}$	-0.002	-0.002	-0.002	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Third Born * $\text{Proximity}_{i,c,t}$	-0.002	-0.002	-0.002	0.003	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Fourth and Above Born * $\text{Proximity}_{i,c,t}$	0.001	0.001	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Obs.Total	977,630					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$ interacted with the son been first (438,793 obs.), second (290,040 obs.), third (146,398 obs.) or fourth and above (102,399 obs.) born; where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1-3), the absolute value of the difference in the Occ. rank between sons and fathers (columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (columns 7-9/10-12). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town , its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5, 6, 8, 9, 11 and 12); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3, 6, 9 and 12). Standard errors clustered at parish level in parenthesis.

Table D.10: Natives vs. Foreigners

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. $cat^{son} \neq$ Occ. cat^{father}			$\mid$ Occ. $rank^{son}-$ Occ. $rank^{father}\mid$		
$\text{Proximity}_{i,c,t}$	0.149	0.116	0.111	1.975	1.489	1.424
	(0.012)	(0.022)	(0.022)	(0.162)	(0.288)	(0.285)
First generation immigrant * $\text{Proximity}_{i,c,t}$	0.001	0.002	-0.008	0.156	0.168	-0.099
	(0.008)	(0.007)	(0.007)	(0.158)	(0.156)	(0.149)
Second generation immigrant * $\text{Proximity}_{i,c,t}$	-0.010	-0.012	-0.011	-0.233	-0.266	-0.251
	(0.007)	(0.007)	(0.007)	(0.149)	(0.148)	(0.148)
	Upward Mobility			Downward Mobility		
$\text{Proximity}_{i,c,t}$	0.067	0.049	0.046	0.036	0.031	0.030
	(0.007)	(0.013)	(0.013)	(0.005)	(0.010)	(0.010)
First generation immigrant * $\text{Proximity}_{i,c,t}$	-0.010	-0.009	-0.017	0.008	0.009	0.005
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Second generation immigrant * $\text{Proximity}_{i,c,t}$	-0.016	-0.017	-0.016	-0.002	-0.003	-0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Obs.Total	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of $\text{Proximity}_{i,c,t}$ interacted with immigrant status been native (915,662 obs.), 1^{st} generation immigrants (24,060 obs.) and 2^{nd} generation immigrants (41,126 obs.); where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (top panel columns 2-3), the absolute value of the difference in the Occ. rank between sons and fathers (top panel columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (bottom panel columns 2-3/4-6). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. The sample include the sample of native sons (columns 1 to 3), 1st generation immigrants (columns 4 to 6), and 2nd generation immigrants (columns 7 to 9). All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3 and 5, 6), and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (columns 3 and 6). Standard errors clustered at parish level in parenthesis.

Table D.11: Locals vs. Outsiders

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. $cat^{son} \neq$ Occ. cat^{father}			$\mid$ Occ. $rank^{son} -$ Occ. $rank^{father} \mid$		
$\text{Proximity}_{i,c,t}$	0.154	0.121	0.114	1.990	1.511	1.414
	(0.012)	(0.022)	(0.022)	(0.163)	(0.288)	(0.284)
$\text{Outsider}^{son}_{i,c,t+1}$ * $\text{Proximity}_{i,c,t}$	-0.024	-0.024	-0.026	-0.162	-0.172	-0.198
	(0.004)	(0.004)	(0.003)	(0.064)	(0.063)	(0.062)
	(7)	(8)	(9)	(10)	(11)	(12)
	Upward Mobility			Downward Mobility		
$\text{Proximity}_{i,c,t}$	0.067	0.049	0.045	0.036	0.030	0.029
	(0.007)	(0.013)	(0.013)	(0.005)	(0.010)	(0.010)
$\text{Outsider}^{son}_{i,c,t+1}$ * $\text{Proximity}_{i,c,t}$	-0.004	-0.004	-0.005	0.001	0.001	0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Obs.Total	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$ interacted with whether the son lives in the same county he was born (745,274 obs.) or not (235,574 obs.); where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1-3), the absolute value of the difference in the Occ. rank between sons and fathers (columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (columns 7-9/10-12). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town , its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5, 6, 8, 9, 11 and 12); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3, 6, 9 and 12). Standard errors clustered at parish level in parenthesis.

Table D.12: Farming occupations

	(1)	(2)	(3)	(4)	(5)	(6)
	Occ. $cat^{son} \neq$ Occ. cat^{father}			$\mid$ Occ. $rank^{son} -$ Occ. $rank^{father} \mid$		
$\text{Proximity}_{i,c,t}$	0.112	0.088	0.084	0.829	0.686	0.611
	(0.013)	(0.023)	(0.022)	(0.149)	(0.270)	(0.267)
Father Farmer * $\text{Proximity}_{i,c,t}$	0.070	0.073	0.070	0.527	0.531	0.509
	(0.003)	(0.003)	(0.003)	(0.052)	(0.052)	(0.051)
	(7)	(8)	(9)	(10)	(11)	(12)
	Upward Mobility			Downward Mobility		
$\text{Proximity}_{i,c,t}$	-0.018	-0.012	-0.015	0.047	0.041	0.040
	(0.006)	(0.011)	(0.011)	(0.006)	(0.010)	(0.010)
Father Farmer * $\text{Proximity}_{i,c,t}$	0.033	0.031	0.030	-0.015	-0.014	-0.015
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Obs.Total	980,848					
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	Yes	Yes	No	Yes	Yes
Household characteristics	No	No	Yes	No	No	Yes

Note: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t}$ interacted with the whether the father was a farmer (220,618 obs.) or not (760,230 obs.); where the Proximity is instrumented by the proximity to the DLCP network in a control function approach. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (columns 1-3), the absolute value of the difference in the Occ. rank between sons and fathers (columns 4-6), and an indicator variable which switches to one if the occupational rank of the son is higher/lower than that of his father and their difference is greater than one standard deviation (columns 7-9/10-12). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier. All regressions include fixed effects for census year and childhood county_t. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town , its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road, port and waterway (columns 2, 3, 5, 6, 8, 9, 11 and 12); and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (column 3, 6, 9 and 12). Standard errors clustered at parish level in parenthesis.