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

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

July 28, 2021

Abstract

Can transport infrastructure promote long-term labour 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 occupation mobility in nineteenth century England and Wales. We create a new dataset of father and son pairs by linking individuals across the full-population censuses of 1851, 1881 and 1911. By geolocating individuals down to the street level, we measure access to the railroad network using the proximity to the nearest train station. To address the non-random access to the railroad network, we create a dynamic hypothetical railroad based solely on geographic cost consideration. We find that sons who grew up one standard deviation (roughly 5 km) closer to the train station were 6 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 labour 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, Louis-Philippe Morin, Andreas Uthemann, 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, and NASMES.

†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 and University of Essex; Research Associate, Centre for Economic Performance, LSE; mmohnen@uottawa.ca

1 Introduction

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". Large transport infrastructure projects have recently been proposed to specifically tackle the rise in inequality. 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, spurring investment in communities that will be better linked to major metropolitan areas, and expanding the middle class.[1] In response, Amtrak presented the "Connects US" plan which not only improves existing lines but also creates new rail corridors which would reach 160 new communities and 20 million more passengers a year.[2]

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 order to empirically assess such long-run effects, one needs a large project that occurred sufficient long ago to track intergeneration mobility. Turning to a historical setting can bring empirical evidence to this open question and shed light on the mechanisms at work.

This paper estimates the causal effect of the access to the railroad network on intergenerational mobility. We exploit the spatial and temporal variation in the expansion of the railroad in nineteenth century England and Wales. According to Rostow (1959), "the introduction of the railroad has been historically the most powerful single initiator of take-offs". It provided new labour opportunities by easing the cost of geographic mobility and changing the landscape of economic activity.

We create a new and unique dataset of close to 1 million father-son pairs for which we observe their occupation, place of residence, and proximity to nearest train station. To do so, we combines several historical data sources. Thanks to the newly digitised full censuses of 1851, 1881 and 1911 for England and Wales (Schürer and Higgs, 2014), we identify individuals across censuses using the linking method proposed by Abramitzky, Mill and Pérez (2019). This allows us to construct a dataset of father-son pairs and measure intergenerational occupation mobility. We geographically locate individuals down to the street level based on their address of residence reported in the census. This dataset permits the analysis of a large and representative sample at a more geographically disaggregated level than was

[1]https://joebiden.com/clean-energy/. The high speed railway linking up London, the Midlands, the North and Scotland (HS2) is expected to cost between £65 and £88 billion and lists as one of its aim to "make Britain better connected, rebalancing the UK economy and bring jobs and investment to the Midlands and North" https://www.hs2.org.uk/why/connectivity/. China has proposed an ambitious high-speed rail (HSR) program that promises to connect all provincial capitals (excluding Lhasa) and large cities with more than half million people by 2030. Some of the justifications behind this expensive project is linking labour markets, facilitating cross-city economic integration and promoting the growth of second-tier cities https://www.nytimes.com/2009/12/23/world/asia/23iht-letter.html.

[2]https://media.amtrak.com/amtrak-connects-us/

previously feasible. By overlaying the digitised railroad network (Alvarez, Bogart, Satchell, Shaw-Taylor and You, 2017), we measure individual's access to the railroad network as the geographical proximity between the place of residence and the nearest train station.

To address the non-random proximity to the railroad network, we create a time series of hypothetical railway maps based solely on geographic cost consideration, ignoring demand-side concerns. This allows us to isolate the portion of the variation that is attributable to exogenous cost considerations and use it as an instrument to estimate the development effects of the transport (Banerjee, Duflo and Qian, 2020; Faber, 2014). The identification strategy exploits the fact that those located along a convenient route were more likely to be 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 6 percentage points more likely to work in a different occupation than their father. They are also 5 percentage points more likely to be upward mobile (i.e. work in an occupation ranked one standard deviation higher than their fathers). These effects appear to be driven by a significant transition out of farming activities and into industrial and commercial activities. The railroad also had distributional consequences. Sons with better access to the railroad network moved away from the middle to the top and bottom 25% of the occupational ranking. It benefitted families at the top and bottom of the occupational ranking. 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 labour market opportunities? We decompose the effect of better access to the railroad network on intergenerational mobility into three channels: the change in the ease of geographic mobility, the change in the relative benefit from moving, and the change in local labour market opportunities. Better connected sons were 10 percentage points more likely to move away from their childhood county. We estimate the return to geographic mobility while accounting for the selection into mobility across households by comparing sons who moved away from their childhood county to their brothers who stayed put (Abramitzky, Boustan and Eriksson, 2012). We find that the railroad decreased the relative benefit from moving. This comes from the fact that the train brought new opportunities locally and/or expanded the labour market of residents thanks to the possibility of commuting. Our decomposition exercise show that the majority of the changes in intergenerational mobility patterns are driven by changes in local labour market opportunities. Local opportunities account for roughly 90% of upward mobility while geographic mobility and the change in the relative benefit from moving only account for 8% and 2% respectively. This imply that when evaluating the effectiveness of transport infrastructure investments, focusing on spatial mobility and disregarding local effects would provide an underestimate of the impact.

Our results provide insights into the long standing debate regarding the approaches to combat inequality and uneven development. On the one hand, "people-based" policies aim to increase the opportunities by targeting directly low-income households (e.g. Moving to Opportunity or Earned Income Tax Credit). On the other hand, "placed-based" strategies aim to increase opportunities by targeting underperforming neighbourhoods (e.g. Empowerment Zone program or European Union Structural Funds). Our results demonstrate that, in nineteenth century England and Wales, the railroad brought opportunities to locally to residents and thus created long-term economic opportunities. This suggests that social mobility issues can be tackled by place-based policies at a local level.

This paper contributes to several strands of the literature. First, there is a large literature on the evaluation of transport infrastructure that has largely focused on aggregate outcomes such as regional trade (Donaldson, 2018; Faber, 2014), agricultural trade and income (Donaldson and Hornbeck, 2016), urbanisation (Baum-Snow, 2007; Duranton and Turner, 2012), regional disparities (Chatterjee and Turnovsky, 2012), and growth (Banerjee et al., 2020). We complement the existing literature by providing individual level responses to these large-scale interventions. Long-run individual outcomes can provide additional insights in the effects of transport infrastructure by revealing heterogeneity in the effects.

We also contribute to the literature documenting intergenerational mobility. The analysis of long run social mobility is complicated by data availability. Researchers have used marriage registrations (Miles, 1999), family histories (Prandy and Bottero, 2000), rare surnames (Clark and Cummins, 2015), and subsamples of census information (Long, 2013). During nineteenth century Britain, Long (2013) show that social mobility is greater than what was previously documented once life-cycle patterns are accounted for. Thanks to newly digitised full censuses, we identify a larger set of father-son pairs than was previously possible. Our sample includes close to 1 million individuals with match rate of 42-49%, thereby allowing us to document intergenerational mobility during the Second Industrial Revolution. Moreover, 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.

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 transportation network and missionary activity are associated with upward mobility. A closely related paper is the one by Perez (2017). He uses the expansion of railroad network in the 19th 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, measures of connectivity and intergenerational mobility, and overall results. The Second Industrial Revolution was an important episode in history that can provide important insights into the nature of growth and social mobility. Moreover, our data allows us to measure connectivity at the individual level. Finally, we show important transitions not only out of farming but into commercial and industrial activities.

The rest of the paper is organised as follows. Section 2 paints the historical background of the railway 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 summarise our findings and conclude in the last section.

2 Background and Data

2.1 Rail Network

Britain was the first industrial and urban society, and the nineteenth century was a time of rapid and dramatic change. The Industrial Revolution marked a period of development with profound social, economic and political change. Treiman (1970) suggests that industrialisation 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 railway was an important driver of this transition.

The main period of railway construction was between 1825 and 1914. The first rail line using steam locomotion was opened in 1825 between Stockton and Darlington in the northern coal mining region. The Liverpool and Manchester railway opened in 1830. It was the first purpose built passenger railway, but also carried freight. There was never a nationwide plan to develop a logical network of railways. Rather, the railway system was promoted by commercial interest and was constructed entirely by private entreprises. During the Railway Mania in the 1840s, England experienced a large railway expansion. This led to a speculative frenzy that reached its peak in 1846 with Parliament authorizing 8,000 miles of lines at a projected cost of £200 million (which was about the same value as the country's annual Gross Domestic Product at that time). By 1870 Britain had about 13,500 miles (21,700 km) of railroad. At the system's greatest extent, in 1914, there was about 20,000 miles (32,000 km) of track, run by 120 competing companies.[3]

[3]The Railway Mania was brought to an end when the government announced closure for depositing

Although the government initially took a laissez faire approach, it was necessary to obtain an Act of Parliament to build a new railway. Almost all railway construction during this period was contested in one form or another. A system of railway hearings was established in the House of Lords, requiring companies to weigh the potential benefit and harm of their proposed schemes.[4] In 1840, the Board of Trade with its Railway Department was created. It was the first government department to assume responsibility for railways. Although the railway was constructed by private firms, the expansion of new routes was also driven by Members of Parliament who wanted to have railway stations in their constituency. Towns were always in competition with their neighbours to attract local trade. They were interested in communication with major cities and other traffic-generating centres, like London.

The railways were part of a steam revolution. Steam power allows the transportation of freight and passenger more quickly and cheaply than before. The railways 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. Railways also facilitated the formation of a international inter-modal transport system by connecting major ports. The first example is the Liverpool-Manchester rail that handled imports of raw cotton and exports of finished cotton goods by linking the the Atlantic port of Liverpool to the textile centre of Manchester. Railways could act as feeders to these ports or as land-bridges between them. 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. Southampton became one of the most successful Atlantic ports, thanks to the London and Southampton Railway.

Steam facilitated speed, and speed attracted passenger traffic as well as freight. The potential of speed encouraged the construction of long-distance inter-urban main lines. Road and canal transport could not compete. By road, the journey between Liverpool and Manchester took four hours and cost 10 shillings inside the coach and 5 shillings outside. By train, 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). The same journey would have taken 20 hours by canal. The cost of canal carriage was 15 shillings a ton, whereas by rail it was 10 shillings a ton.

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

---

schemes. The Railway Grouping of 1923 terminated much of the inter-company competition that characterized the 1825-1914 period.

[4] `https://www.mtholyoke.edu/courses/rschwart/ind_rev/rs/denault.htm`

Figure 1: Railway Network, 1851-1881

(a) 1851

(b) 1881

2.2 Intergenerational Mobility

2.2.1 Linking Individuals Across Censuses

Our aim is to relate intergenerational mobility and access to the railroad network at the individual level. For this purpose, we combine several historical sources to create a new and unique dataset. We first use the 100% sample of England and Wales census from 1851 to 1911 developed by the I-CeM project. The data collection contains records for more than 35 million households and over 180 million individuals. The census contains the full address of individuals (house number or name, name of street, avenue or road, civil parish and county of residence). In addition to geographic variables, the census also provides a wider range of sociodemographic information: age, gender, place of birth, marital status, number of children, number of servants and family structure as well as information on occupation defined as that in which the individual was principally engaged on the day on which the census was taken. The only economic outcome available in our data is self-reported occupation. There are over 400 occupations.[5]

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. As women may have

[5]There are 1.4 million occupational strings in the 1881 census.

changed their surname due to marriage, we focus on men. Records were only compared in the linking process if they had an exact match on parish or county of birth. Age was allowed to be up to two years higher or lower than would be expected, while first and last names were allowed to have a Jaro-Winkler distance no larger than 0.1 (Jaro, 1989). Individuals are 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.[6] As the censuses record 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 family name between the father-son should have a Jaro-Winkler distance no larger than 0.12 to guarantee that the father-son pair are from the same family. We also restrict sons to be between 40 and 52 and fathers to be between 20 and 65 years old. This restriction guarantees that we are looking at men during their working years.

We link 652,192 father-sons pairs in 1851-1881 and 1,227,324 in 1881-1911. This represents approximatively 42-49% of the population. As a point of comparison, match rates in other studies are between 7-42% (see Table A.3).[7] Section A.6 presents descriptives of the linked sample, showing that it is a representative sample of the full census. In particular Table A.5 shows that the role of the railroad network in explaining the share of linked individuals is limited.[8]

2.2.2 Intergenerational Occupational Mobility

Linking individuals across censuses allows us to observe an individual's occupation as an adult (aged between 40 and 52) and his father's occupation during his youth (when the individuals was aged between 10 and 22 and his father was aged between 40 and 52). 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

[6] The details of the matching procedure and the representativeness are described in the Appendix A.2.

[7] The reason behind our higher match rate is the fact, 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. Milner (2020) matches men in the England and Wales census from 1861 to 1881 and 1881 to 1901 with a very high match rate of 37 and 42%, respectively. Guerra and Mohnen (2020) match the census 1851 to 1881 for London only.

[8] 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 distance to the train station is somehow related to the survival probability.

done in historical setting (Boberg-Fazlic, Sharp et al., 2013; Clark and Cummins, 2015; Ferrie, 2005; Long and Ferrie, 2013; Olivetti and Paserman, 2015). Although earnings is the measure most commonly used in economics, there are advantages to using occupation. One advantage is that occupations are more stable to transitory income shocks over the life cycle than income. Moreover, occupation captures more dimensions of an individual's experience that may be related to interpretations of social mobility such as prestige in the community, autonomy in the workplace, manual versus non-manual labour, and place of work.

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).[9] There are 359 unique HISCAM scores, ranging from 28 to 99, and higher scores indicate a more advantageous position in society. We use the variation as a continuous variable which we standardise in our main specifications. We also define two indicator variables "upward mobility" and "downward mobility". The former (latter) switches from zero to one if the son' occupation has higher (lower) HISCAM score than the occupation of his father and the difference in HISCAM scores between father and son is higher than one standard deviation of the son's distribution.[10] Since we are interested in occupation mobility between father and son we employ a HISCAM ranking that is constant over time. However, the status or socio-economic position of an occupation may vary over time especially with the transition to industrialisation (e.g. being a farmer in 1851 may not reflect the same prestige as being a farmer in 1881 (Xie and Killewald, 2013)). If we were interested in the mobility in terms of socio-economic status, we would need a ranking in which an occupation's position can change over time. We perform our analysis using such a ranking as a robustness check (see Figure D.2).

Occupational ranking captures the potential for higher mean earnings for each occupation and the potential for occupational upgrading. We therefore also use four occupation classifications to detect nonlinearities through occupation transitions. The first classification is based on the Historical International Standard Classification of Occupation (HISCO). HISCO is not a class or status scheme but rather a classification by economic sector or workplace tasks. It groups occupations into major groups and further divides into minor groups.

[9]Some people refer to the structure measured by HIS-CAM/CAMSIS scales as a structure of 'status', 'prestige', 'socio-economic position', or 'class'. The 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 score is rescaled to a mean of 50 and a standard deviation of 10, with higher scores indicating a more advantageous position in society. Occupational combinations of parents and their adult children is not a problem when examining intergenerational mobility in occupation since occupations do not perfectly predict intergenerational mobility. See HISCAM scale at `www.camsis.stir.ac.uk/hiscam`. It was developed to facilitate the analysis of data coded to the Historical International Standard Classification of Occupations (HISCO).

[10]More formally, let H^s be the HISCAM score of the son, with standard deviation $\sigma^s = \sqrt{Var(H^s)}$. We define a son as upward mobile if $H^s > H^f$ and $|H^s - H^f| > \sigma^s$. We also examine the robustness of our results to different definitions of upward and downward mobility (see Figure D.2).

At the finest level of detail it includes 1881 occupational categories. In our paper, we present seven major groups: professional, managerial, clerical, sales, services, farm and labourer.[11] The second classification recodes historical occupations into a class scheme, a status scale and a division into economic sectors. These include professional, industrial, commercial, domestic, and agriculture (Woollard, 1998).[12] We refer this classification as Woollard occupation categories. The third classification is the HISCLASS which categorises 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.[13] Finally, we create a binary variable for the literacy of an individual based on Armstrong (1972)'s measure of the literacy requirement of each occupation.

2.2.3 Geolocating Individuals

To geo–reference individuals, we use the Great Britain addresses (GB1900) (Southall, Aucott, Fleet, Pert and Stoner, 2017), and the digitised parish and county boundaries provided by the UK Data Service (Satchell, Kitson, Newton, Shaw-Taylor and Wrigley, 2017). We locate individuals down to the street level within their parish. For this we perform a string matching on address of residence (street name and parish) from the census and the street points within each parish. Any measurement error in the location of individual is limited to the parish boundaries. To remove any measurement error, we also locate individuals to the centroid of their parish as a robustness check. In the baseline specification, we measure individual access to the railroad network based on the shortest straight line distance between their place of residence and the nearest train station.

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 a HISCAM of 50 and 49 respectively. HISCAM between 49 and 50 include a broad range of occupations such as farmer, labourers, professionals and services. 18% of sons experience upward mobility

[11] "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 labourers and servants, "Labourers" include for instance coal miners, carpenter, and painters.

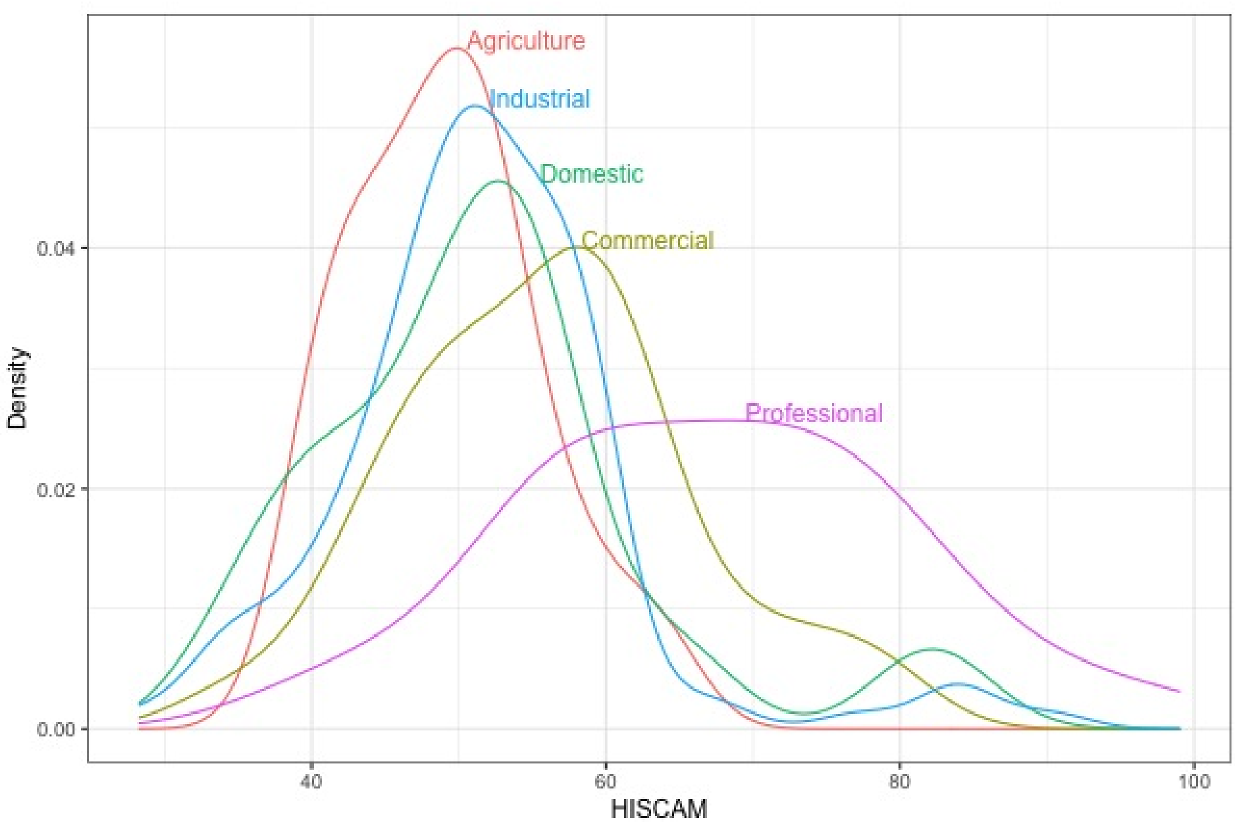
[12] "Professional" include schoolmasters/teachers, police, postmen, solicitors, "Industrial" include general labourers, coal miners, carpenters, "Commercial" comprise of commercial or business clerks, "Domestic" comprise of domestic and servants, and "Agriculture" includes all agriculture-related activities.

[13] "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 labourers, coal miners, or drivers.

while 15% experience downward mobility, where sons are considered upward (downward) mobile if they have a higher (lower) HISCAM occupation rank than their father and this difference is larger than 1 standard deviation. 32% of sons move away from the county they grew up in. Moreover, sons move far away from where they grew up. On average they move 98km further away.

Figure 2 presents the distribution of the HISCAM occupation ranking by Woollard occupation categories. There was 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 high HISCAM score and agriculture occupations having on average a low HISCAM scores. Nevertheless, occupational ranking and occupation categories contain different information. Within each occupation category, there is a range of occupational ranking. For instance, within professional occupations monks have a high occupational ranking while soldiers have a low occupational ranking. It is therefore important to examine both the occupational ranking and occupation categories when exploring intergenerational mobility. Figure A.2 in the Appendix presents the density plots for HISCO and HISCLASS occupation categories.

Figure 2: HISCAM distribution by Woollard occupation category, 1851-1911



Note: This plots displays the density of HISCAM by Woollard occupation categories

The correlation in HISCAM occupational rank between fathers and sons is 0.28. Table

Table 1: Descriptive Table

	Mean	St. Dev.	Min.	Median	P75	Max.
			A. SONS			
Age	44.64	3.51	40	44	47	52
Foreign-born	0.02	0.15	0	0	0	1
Urban resident	0.39	0.49	0	0	1	1
Literate	0.32	0.47	0.00	0.00	1.00	1.00
HISCAM occupation rank	50.10	10.11	28.28	50.81	57.20	99.00
$HISCO^{son} \neq HISCO^{father}$	0.80	0.40	0	1	1	1
$\|HISCAM^{son} - HISCAM^{father}\|$	8.03	8.39	0	5.9	12.6	71
Upward mobility	0.18	0.39	0	0	0	1
Downward mobility	0.15	0.36	0	0	0	1
Dist. to nearest train station (in km)	3.25	5.41	0.005	1.49	3.61	106.57
County mover	0.32	0.47	0	0	1	1
Dist. moved | county mover	98.56	96.98	0.02	68.74	144.81	633.24
			B. FATHERS			
Age	46.67	7.61	20	46	52	65
Foreign-born	0.05	0.22	0	0	0	1
Urban resident	0.39	0.49	0	0	1	1
Household size	6.76	2.15	0	7	8	45
Number of sons	4.63	2.09	0	5	6	17
Number of servants	0.17	0.65	0	0	0	54
HISCAM occupation rank	49.41	9.12	28.28	50.95	53.50	99.00
Literate	0.31	0.46	0.00	0.00	1.00	1.00
			C. COUNTY			
Number of father-son pairs	17,622.58	20,678.73	498	12,590	21,604.5	110,755
Area (km^2)	2,738.97	1,605.23	1.52	2,212.82	3,665.55	7,135.78
Population	172,660.70	245,294.20	6,633	96,023	187,051.8	1,448,853
Avg. HISCAM	49.60	1.91	44.00	50.14	50.77	53.37
Avg. dist. to train station (in km)	5.62	5.39	0.81	3.88	5.81	26.32

Note: The sample consists of 969,242 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).

2 provides a cross-classification of sons and fathers' occupations. We distinguish between son's within 5km of the train station ("connected") and those growing up further away

("non-connected"). Regardless of connectedness, sons tended to follow their father's occupation as the larger percentage is found along the diagonal. Nevertheless, there is still important features worth highlighting. First, sons whose father were lower skilled workers appeared to be the least mobile. This is largely due to the excessive size of this category. Second, upward mobility was substantial, but not across the entire distribution of classes. For example, between 30 and 34% of sons whose fathers were lower skilled were upward mobile (becoming skilled workers or managers), while between 12 and 16% of sons of skilled workers were upward mobile. Third, connected sons experienced slightly greater mobility than non-connected sons. The total mobility, as measured by the share of individuals off the main diagonal, is 51% for connected sons and 50% for non-connected sons. 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. Fourth, connected sons whose fathers were in top occupations were more likely to stay in top occupations than non-connected. In contrast, connected sons from fathers who were at the bottom are less likely to stay at the bottom than non-connected sons. Finally, farmers constituted a larger share of the sample among connected sons. These sons experienced the most upward mobility. Importantly, the magnitude of intergenerational mobility and patterns observed here are similar to those found in previous studies.[14]

A new feature of our dataset is the ability to geographically locating individuals. Figure 3 shows the average connectivity by county. Most individuals lived within 5 to 10km away from the nearest train station. Residents of Wales and Cornwall were the least connected to the railroad network. They lived between 10 and 27km from the nearest train station. In contrast, residents of Manchester, Liverpool and Birmingham lived within 2.5km of the nearest train station.

Figure 4 reveals striking spatial patterns in intergenerational mobility. In the first four subfigures we see intergenerational mobility measures plotting across counties. We see that in places of opportunity such as London and many coastal towns sons tended to have higher intergenerational mobility. Sons who grew up in the south of England (e.g. Devon, Somerset, Dorset) were less likely to follow the occupation of their father than sons who grew up in Cornwall, Wales and the north of England (e.g. Durham). However, these patterns do not necessary match the average distance in occupation ranking between fathers and sons. Sons from the northern counties of (e.g. Northumberland) for instance were more likely to follow the occupation of their father than those in the east of England (e.g. Norfolk), but they show the opposite social mobility pattern in terms of distance in occupational ranking. This highlights the importance of looking at both the intensive and extensive margins. Finally, when looking at the probability of upward and downward mobility we also see large variation

[14] Long (2013) measures the occupation intergenerational mobility for 1851-1881 and 1881-1901. Fathers' occupations are observed when sons were age 10-19. For 1851-1881 (1881-1901), he finds that the rate of total mobility is 50.1% (48.3%), the rate of upward mobility is 26.8% (26.8%), and the rate of downward mobility is 23.3% (21.5%). Miles (1999) found a total mobility of 34.8% and upward mobility is 17.7% using a sample of marriage registries from 1859-1874. Jantti, Bratsberg, Roed, Raaum, Naylor, Osterbacka, Bjorklund and Eriksson (2006) estimates the correlation coefficient father-son pairs to be 0.198 using the National Child Development Study in the UK in 1974.

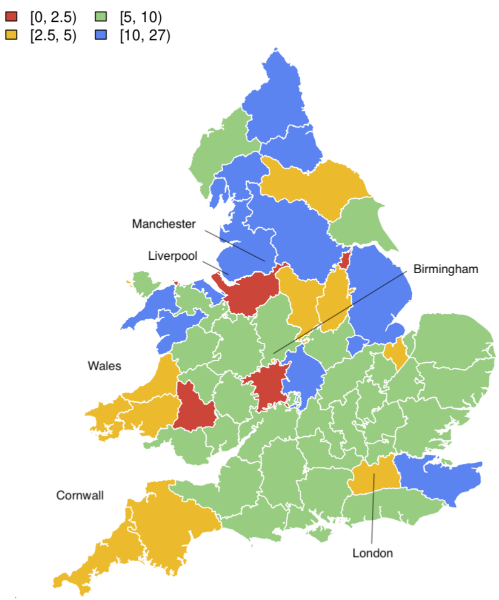
Table 2: Mobility Matrix for Connected Sons

	Connected Father				
Son	Manager	Skilled Workers	Lower Skilled	Farmers	Total
Manager	**0.402**	0.161	0.147	0.122	
	{37,385}	{30,935}	{47,097}	{17,065}	132,482
Skilled Workers	0.192	**0.415**	0.189	0.127	
	{17,858}	**{79,630}**	{60,692}	{17,709}	175,889
Lower Skilled	0.348	0.379	**0.609**	0.358	
	{32,341}	{72,772}	**{195,105}**	{50,054}	350,272
Farmers	0.057	0.044	0.055	**0.393**	
	{5,304}	{8,521}	{17,565}	**{54,974}**	86,364
Total	92,888	191,858	320,459	139,802	745,007
	Non-connected Father				
Son	Manager	Skilled Workers	Lower Skilled	Farmers	Total
Manager	**0.336**	0.120	0.117	0.099	
	{3,553}	{3,733}	{4,228}	{8,001}	19,515
Skilled Workers	0.193	**0.483**	0.178	0.104	
	{2,035}	**{15,007}**	{6,460}	{8,403}	31,905
Lower Skilled	0.290	0.282	**0.534**	0.283	
	{3,066}	{8,762}	**{19,310}**	{22,867}	54,005
Farmers	0.181	0.114	0.171	**0.513**	
	{1,914}	{3,539}	{6,196}	**{41,447}**	53,096
Total	10,568	31,041	36,194	80,718	158,521

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 51% for connected sons and 50% for non-connected sons. Occupation classification is based on HISCLASS.

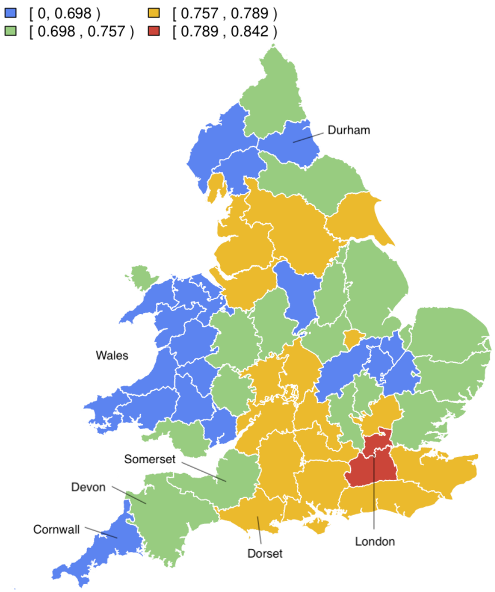
across England and Wales. In some places there was both a high upward and low downward mobility such as Lancashire and Manchester, both of which were specialized in manufacturing. Other places experienced low upward and high downward mobility. This was the case of Nottingham, famous for its textile industry and its slums. Places such as London, Devon or the south of Wales experienced both high upward and downward mobility. Finally, some places experienced both low upward and downward mobility. This was the case in East Anglia where there were many wealthy estate owners.

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

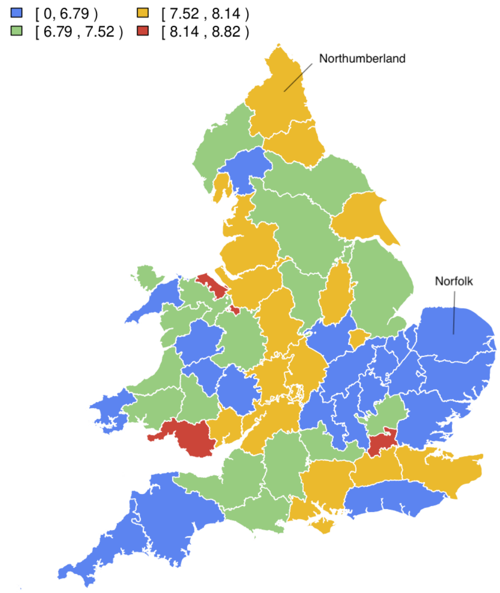


Note: This figure presents the quantile of the average distance between place of residence and the nearest train station by county.

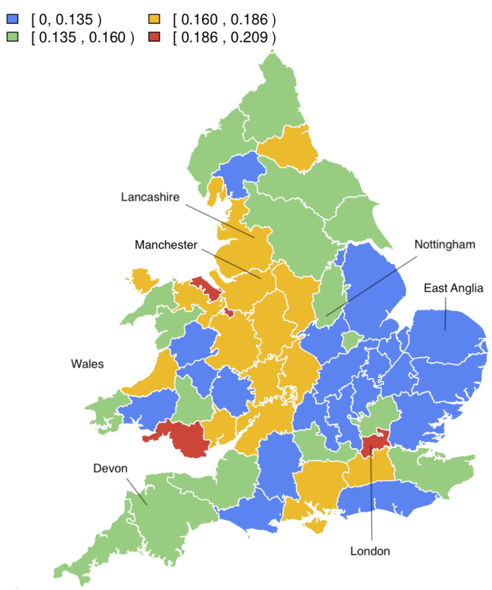
Figure 4: Mobility patterns by county at $t-1$



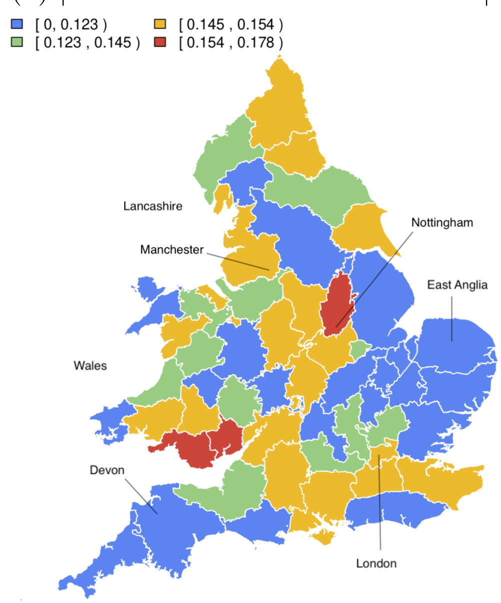
(a) $\Pr(HISCO^{son} \neq HISCO^{father})$



(b) $|HISCAM^{son} - HISCAM^{father}|$



(c) Pr (Upward mobility)



(d) Pr (Downward mobility)

Note: This figure presents the quantile of four intergenerational mobility measures by county where sons grew up.

4 Empirical Strategy

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

$$f(Rank_{i,c,t}^{son}, Rank_{i,c,t-1}^{father}) = \alpha_1 Proximity_{i,c,t-1} + \alpha_2 X_{i,c,t-1} + \gamma_t + \rho_c + \epsilon_{i,c,t-1} \tag{1}$$

where i, c, and t index family, county of residence at time $t-1$ and census year respectively. The dependent variable can take various forms: (1) an indicator variable equal to one if the son works in a different HISCO occupation than his father, (2) the absolute difference between the HISCAM scores of the father and son, (3) a dummy variable equal to one if the son's HISCAM score 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) a dummy variable equal to one if the son's HISCAM score 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 access to the railroad network, $Proximity_{i,c,t-1}$, as the standardised proximity (i.e. negative standardised distance in kilometres), measured as a straight line between the place of residence and the nearest train station at $t-1$. 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 railway 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, this variable is measured using indicators equal to one if the son grew up within 5, 10 and 15km of a train station or whether the parish of residence at that time had a train station within its boundaries. We measure connectedness during youth at $t-1$ when the sons lived with their fathers. In our setting, sons are between 10 and 22 years old. At this time, sons are still living with their father and have not yet become a head of household themselves.

Finally, we include a vector of control variables $X_{i,c,t-1}$ which we discuss below. 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 labour 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_1 captures the effect of growing up one standard deviation closer to the nearest train station on intergenerational mobility. There could be serial correlation in the error term $\epsilon_{i,c,t-1}$. We therefore cluster standard errors at the level of the parish of residence measured at time $t-1$.

4.1 Dynamic least cost railroad network

Estimating equation 1 by OLS would imply that, conditional on controls, year and county, the proximity to the railroads would have to be exogenous. This would be the case if the railroad lines and train stations were randomly built across England and Wales.

Given the high cost and potential large benefits of infrastructure investments, the placement of new railway lines was most likely correlated with the demand for trade, migration and/or important local labour markets. This raises the concern that connected locations were more or less likely to grow in the future, regardless of the railroad construction. If railroads were built between cities that were expected to have high economic growth, then economic growth and not railroads may be the driving force for observed differences in mobility patterns. It may also be the case that favourable labour market shocks happened to hit locations that were recently connected by the rail network, and this is what drives mobility. In such cases, the OLS would overestimate the effect of being connected. Alternatively, if the railroads targeted struggling areas with limited social mobility, the OLS would underestimate the effects.

Moreover, the place of residence within a county is unlikely to be random. If wealthier families, that experience higher upward mobility patterns, were more likely to live closer to town centres and the train stations were generally located close to town centres, then the OLS estimates would overestimate our effect. In contrast, if poorer families that experienced limited social mobility lived close to train stations the OLS would underestimate our results.[15]

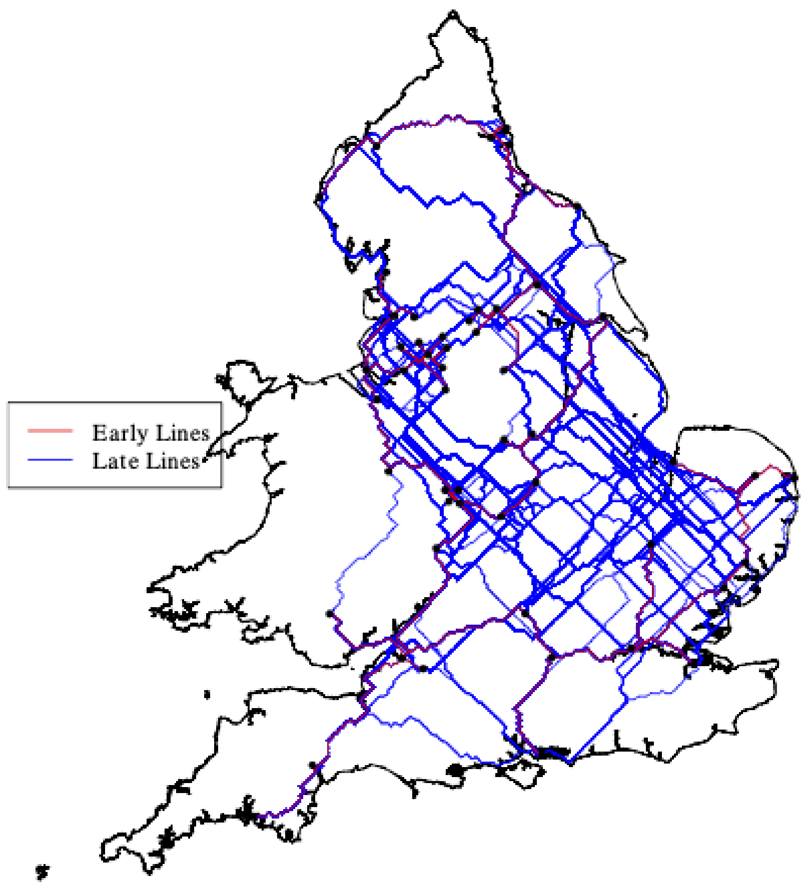
To address the endogenous proximity to the train station, we use the "inconsequential place IV approach" (Alvarez et al., 2017; Banerjee et al., 2020; Duflo and Pande, 2007; Faber, 2014; Lipscomb, Mobarak and Barham, 2013). We construct a hypothetical railroad network showing how the railway would have evolved had planners only considered geographic cost and ignored demand-side concerns. We proceed in three steps. In the first step, we identify major towns in 1801.[16] By taking population at that time, we avoid any possible confounder related to population growth induced by the railroad. In the second step, 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 towns is determined by minimising the slope cost of all the cells the path crosses. Each cell with slope s has a crossing cost $1 + \left(\frac{s}{S}\right)^2$, with S being a slope threshold that we set at the median slope of the observed network (Herzog, 2013). In a final step, we distinguish between rail lines that were likely to be constructed earlier than others. For this, we compute the least cost network connecting all major 1801 towns as the early projected line. In doing this, we give higher weight to those

[15] Figures B.1 and B.2 in the Appendix show the relationship between the distance to the nearest train station and the wealth of a family, as measured by the father's HISCAM and the number of servants respectively. We see that fathers living within 5km to the nearest train station belong to the full range of HISCAM ranking and the number of servants, while those living further away tend to work in lower ranked occupations and have fewer servants. The colour highlights the intergenerational mobility patterns. Upward and downward mobilities occurs at all distances. Naturally upward mobility is correlated to the father's HISCAM and only possible if the father belongs to the lower HISCAM ranking.

[16] Within all towns in 1801, we consider a town a major town if it belongs to the top 10% of the population distribution. This represents towns with at least 9,172 inhabitants in 1801. There is a total of 53 towns in the top 10%.

network edges connecting larger towns.[17] The remaining least cost path network connecting all pairs of major towns is the late projected line. 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.

Figure 5: Projected Railroad Lines



Note: The green crosses 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.

While our $Proximity_{i,c,t-1}$ is defined as the proximity (in km) between the place of residence and the nearest train station, the instrument is defined as the proximity between the place of residence and the DLCP network.[18] Therefore, the first stage equation is defined

[17] For an edge connecting towns p and q the cost of implementing it is:

$$\text{edge cost}(p, q) = \text{slope cost}(p, q) + \left(\frac{pop_p + pop_q}{\max_{k,l \in town: k \neq l}(pop_k + pop_l)}\right)^{-1}$$

where the slope cost is obtained by aggregating the slope cost of each cell that the (p, q) edge crosses.

[18] Figure B.3 illustrates the instrument.

as:

$$Proximity_{i,c,t-1} = \beta_1(Proximity\ to\ DLCP)_{i,c,t-1} + \beta_2 X_{i,c,t-1} + \gamma_t + \rho_c + \eta_{i,c,t} \qquad (2)$$

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

The instrument based on the DLCP railroad network addresses the endogeneity in the proximity to the nearest train station stemming from the construction decision of the railroad and the location decision of families. It isolates the portion of the variation in the expansion of the railroad that is attributable to exogenous cost considerations. Moreover, the DLCP railroad 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 and town centres. Location decision may be based on local amenities but is unlikely to be correlated with the relative path to other town centres 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.[19]

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 hypothetical 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 towns further away. Proximity to major economic centres might also be correlated with economic characteristics of the towns which also affect growth trajectories. This in turn would have a direct effect on the economic opportunities of town residents. We address this concern by including the distance to the closest 1801 town, their 1801 populations and the 1801 population in the surrounding area.[20] These variables proxy for the historical importance of the place of residence within a county as a traffic junction and a likely stop for the railroad.

The DLCP railroad 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 the initial travel routes and not the new railroads. We control for the proximity to historical places of trade as proxied by ancient ports (Alvarez-Palau and Dunn, 2019) and Roman Roads (McCormick, Huang, Zambotti and Lavash, 2013).

[19]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 of their location is convenient close to the DLCP network but would not have been close otherwise. In the robustness check, we compute the causal response weighting function.

[20]The 1801 population in the surrounding area is measured using the following equation: $\sum_{p \neq q} Pop_p / D_{p,q}$ where Pop_p is the standardised population of parish p and $D_{p,q}$ is the standardised distance between the centroids of parishes p and q. It is measured at the parish where the individual was living in $t-1$.

The place of residence may be correlated with underlying intergenerational mobility patterns. To the extent that the initial wealth of a family determines both the place of residence and the experienced intergenerational mobility, the distance to the train station may be picking up family characteristics. We 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 individuals residing along the DLCP railroad network affects experience changes in 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 3 we see a positive and statistically significant correlation between the proximity to the rail station and the proximity to the hypothetical railroad 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.

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 4 presents the causal effect of being one standard deviation (approximately 5km or one hour's walk) 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 moved up in the occupational ranking (row 1). They were not only less tied to their father's occupation (row 2) but also moved further away from the occupation ranking of their father (row 3). Moreover, they experienced upward and downward mobility relative to their father (rows 4 and 5 respectively). 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 occupation ranking 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 includes all control variables in addition to county and census year fixed effects. This is our preferred specification for the remainder of our paper. Sons who grew up one standard deviation (approximately 5km or one hour's walk) closer to the train station were 6 percentage points more likely to work

Table 3: First stage regressions

	(1)	(2)	(3)
Dep. var.:		$\text{Proximity}_{i,c,t-1}$	
Proximity to DLCP $\text{network}_{i,c,t-1}$	0.016***	0.014***	0.014***
	(0.001)	(0.001)	(0.001)
Obs.		969,242	
R^2	0.401	0.440	0.440
SW-F	110.738	24.118	14.454
F-Stat	110.738	144.707	144.536
Year FE	Yes	Yes	Yes
County_{t-1} 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 the standardised proximity between the childhood residence and the nearest train station and the independent variable in the standardised negative distance between the childhood residence and the nearest railroad line from the DLCP network. All regressions include fixed effects for census year and childhood county_{t-1}. 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 and port (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). SW-F reports the F-stat from Sanderson and Windmeijer (2016). Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

in a different occupation than their father. They were also 5 percentage points more likely to be upward mobile. To illustrate these effects, we look at a concrete example from our dataset. Two sons whose fathers were farmers (HISCAM = 39.58), one grew up 5.06km from the nearest train station and became a manager (HISCAM = 84.75), while the other grew up 20.20km from the nearest train station and became a labourer (HISCAM = 53.04).[21]

[21] The effects of being better connected to the railroad network are likely to be non-linear with sons living within a certain distance benefitting from being connected and those beyond a certain distance no longer being connected. 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 (see figures D.4 and D.5).

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 placement of train stations. OLS regressions underestimate the gains from connectivity. This is consistent with the railroad locations targeting areas with limited intergenerational mobility and particularly upward mobility. The OLS estimates could also be biased due to classical measurement error in the railroad access corrected by the IV estimate. Finally, 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 but would not have been so close otherwise.[22]

[22] 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.

Table 4: The effect of railroad connection on intergenerational mobility

	(1)	(2)	(3)	(4)	(5)	(6)
		OLS			IV	
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.020***	0.011***	0.010***	0.089***	0.065***	0.062***
	(0.001)	(0.001)	(0.001)	(0.010)	(0.010)	(0.010)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	0.370***	0.238***	0.228***	1.249***	1.102***	1.057***
	(0.023)	(0.022)	(0.022)	(0.132)	(0.147)	(0.144)
Upward Mobility	0.014***	0.009***	0.008***	0.051***	0.050***	0.049***
	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)	(0.006)
Downward Mobility	0.005***	0.003***	0.003***	0.012***	0.007	0.006
	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)	(0.005)
Obs.	969,242					
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County_{t-1} 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 standardised $\text{Proximity}_{i,c,t-1}$ 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 HISCAM 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 childhood county_{t-1}. 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 and port (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 in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

5.3 Effects by occupation categories

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 the transition between occupations.

$$\Pr(Occ^{son}_{i,c,t} = k | Occ^{father}_{i,c,t-1} = m) = \alpha_1 Proximity_{i,c,t-1} + \alpha_2 X_{i,c,t-1} + \gamma_t + \rho_c + \epsilon_{i,c,t} \quad (3)$$

where $\Pr(Occ^{son}_{i,c,t} = k | Occ^{father}_{i,c,t-1} = m)$ is the probability that a son with a father in occupation category m would work in occupational category k. Just as in Equation 1, $proximity_{i,c,t-1}$ is defined as the standardised proximity between the place of residence and the nearest train station at $t-1$. The control variables are the same as in the previous most complete specification.

Table 5 presents the results from Equation 3 for the HISCO occupation classification. It reveals some interesting patterns. First, sons who grew up closer to the railroad network moved out of farming occupations regardless of their father's occupation. They were also more likely to work as labourers. This is consistent with the railroad reinforcing the effects of the Industrial Revolution which involved a decline in the proportion of agricultural workers and an increase in the prevalence of industrial and commercial activities. Second, better connected sons were also significantly more likely to move into professional occupations. Third, we see a large variation in the effect of being better connected to the railroad network on the transition within and across occupations. For instance, better access to the train station for sons of salesmen significantly increased their probability of becoming a labourer or a clerk, but decreased their probability of staying in sales. Sons who grew up closer to the train station whose father worked in clerical occupations were more likely to become professionals by 7 percentage points. In contrast, better connected sons of managers saw an increased chance of becoming labourers by 15 percentage points.

To provide additional insight into the transition, we also present transitions between Woollard occupations in Tables C.1 in the Appendix. Again, we see a large and significant transition out of farming activities. Conditional on the father working in agriculture, better access to the railroad increased the probability of working in a domestic activity by 3 percentage points and industrial activities by 10 percentage points. Moreover, connection to the railroad significantly increases the probability of working in commercial and industrial occupations. The Industrial Revolution was a period of important transitions in the production processes which had consequences for the overall occupation structure. We explore the effect of railroad access on occupations that grew or decline between 1951 and 1911.[23] In Table C.6 in the Appendix, 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.

[23]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.2 in the Appendix presents examples of occupations with the highest and lower growth.

Table 5: 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	All
Professional	0.002	0.114***	0.073***	0.009	−0.010	0.004	0.009***	0.012***
	(0.028)	(0.042)	(0.019)	(0.010)	(0.009)	(0.003)	(0.002)	(0.003)
Managerial	−0.015	0.034	−0.002	0.001	0.025***	−0.002	0.001	−0.0001
	(0.014)	(0.037)	(0.018)	(0.007)	(0.008)	(0.003)	(0.002)	(0.001)
Clerical	0.041**	0.016	−0.002	0.026***	−0.001	0.005*	0.003	0.007**
	(0.018)	(0.036)	(0.031)	(0.010)	(0.013)	(0.003)	(0.003)	(0.003)
Sales	0.056**	−0.015	−0.022	−0.063***	0.001	0.001	−0.005	0.004
	(0.023)	(0.047)	(0.031)	(0.024)	(0.015)	(0.005)	(0.005)	(0.005)
Services	−0.034	−0.072*	−0.026	0.009	0.025	0.002	0.012**	0.007*
	(0.021)	(0.038)	(0.024)	(0.008)	(0.018)	(0.004)	(0.005)	(0.004)
Farm	−0.025	−0.230***	−0.041**	−0.054***	−0.063***	−0.111***	−0.051***	−0.135***
	(0.020)	(0.078)	(0.018)	(0.014)	(0.017)	(0.017)	(0.008)	(0.015)
Labourer	−0.023	0.153**	0.020	0.074***	0.022	0.100***	0.031***	0.105***
	(0.037)	(0.075)	(0.044)	(0.024)	(0.025)	(0.015)	(0.010)	(0.012)
Obs.	22,269	15,285	21,357	75,318	34,226	226,466	574,321	969,242

Notes: Each coefficient represents the coefficient of the standardised Proximity$_{i,c,t-1}$ instrumented by the proximity to the DLCP network. 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). Additional sample restriction are that the fathers work as "professionals" (column 1), "managers" (column 2), "clerical" (column 3), "sales" (column 4), "services" (column 5), "farm" (column 6) and "labourer" (column 7). All regressions include census year and childhood county$_{t-1}$ fixed effects, 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. *p<0.1; **p<0.05; ***p<0.01

It has been shown that the railroad lead to higher school enrolment and increase skill premia in the local labour market (Atack, Margo and Perlman, 2012; Michaels, 2008). All else equal, such educational investments will allow these sons to work in higher-ranked occupations than their father. Unfortunately, the historical censuses do not have information on education level. Instead we measure skill and literacy based on occupations. In Table C.2 in the Appendix, we see that better connected sons are 8 percentage points more likely to be literate and 4 percentage points more likely to work in a high-skilled occupation.[24] When conditioning on fathers being illiterate or unskilled, we see that sons are upward mobile in terms of these skills. Table C.3 disaggregates occupations by skill levels using the HISCLASS ranking and presents the transitions between these occupations. As observed previously, we see that there is a general movement out of farming. We also see a significant transition induced by the railroad from lower skilled workers to skilled workers and vice versa. Better connected sons of lower skilled workers moving up to skilled worker while sons of skilled

[24] Skill level is defined as an indicator variable equal to one if the HISCLASS occupational ranking is "manager", "skilled worker" or "lower skilled".

workers moving down to lower skilled worker. In sum, our results show that the railroad network improved the skill and literacy attainment of connected children.

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 the distributional effect of the expansion of the railroad network, we divide the HISCAM occupational ranking of fathers and sons into “Upper”, “Middle” and “Lower”, where “Upper” and “Lower” represent the top and bottom 25% of the distribution respectively. We estimate the following equation

$$Pr(Rank^{son}_{i,c,t} = Q | Rank^{father}_{i,c,t-1} = P) = \alpha_1 Proximity_{i,c,t-1} + \alpha_2 X_{i,c,t-1} + \gamma_t + \rho_c + \epsilon_{i,c,t} \quad (4)$$

where Q and P are to upper, middle and lower groupings. We refer to them as “classes”.

Table 6 presents the effect of being closer to the nearest train station on the conditional probability of being in a certain class. We see that the benefits from the railroad network were not uniform across classes. Sons from upper class families benefitted the most from better access to the railroad network. For them, growing up next to the train station as oppose to one hour’s walk meant that they had a significant 15 percentage points higher probability of staying in the upper class. Sons coming from the middle class families benefitted the least. Being closer to the train station represented a significant 4 and 5 percentage points increase in the probability of moving down and up in class respectively. Finally, for sons from lower class families, access to the railroad network significantly improved their chance of moving to the upper class. These results suggest that the railroad network shifted the distribution to both tails of the distribution with a shrinking middle class.[25]

[25] In the Appendix, we further investigate the differential effects by family background. We condition on the father being in a white or blue-collar occupations in C.4. Sons whose fathers in a blue collar occupation experience larger benefits from better access to the railroad network. In Figure C.1, we disaggregate the distribution of son and fathers’ HISCAM by the percentiles. 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. 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.

Table 6: Distributional Consequences

	(1)	(2)	(3)	(4)
	Father			
	Bottom	Middle	Top	Any
Bottom	0.011	0.044***	−0.011	0.035**
	(0.022)	(0.010)	(0.010)	(0.014)
Middle	−0.061***	−0.095***	−0.137***	−0.116***
	(0.019)	(0.012)	(0.017)	(0.014)
Top	0.050***	0.051***	0.149***	0.081***
	(0.012)	(0.009)	(0.018)	(0.011)
Obs.	235,909	446,906	286,427	969,242

Notes: Each coefficient represents the coefficient of Proximity$_{i,c,t-1}$ instrumented by the proximity to the DLCP network. The dependent variable is an indicator equal to one if the son is at the bottom 25% (row 1), middle (row 2), or top 25% (row 3) of the HISCAM distribution. Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier (column 4). Additional sample restriction are that the fathers is at the bottom 25% (column 1), middle (column 2) or top 25% (column 3) of the HISCAM distribution. All regression include census year and childhood county$_{t-1}$ fixed effects, 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. *p<0.1; **p<0.05; ***p<0.01

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 remain when using alternative measures of connectedness, and intergenerational mobility, and empirical specification. In Figure D.1, instead of measuring the proximity to the nearest train station, we use the distance 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, instead of defining upward (downward) mobility as an indicator variable taking the value one if the son has a higher (lower) HISCAM occupational ranking than his father and the difference is at least one standard deviation, we use 0.5, 1.5 and 2 standard deviations. Our baseline HISCAM occupation ranking is consistent over time. However, we also consider an alternative time-varying occupational ranking. 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 effect in Figure D.3.

Second, we explore potential measurement errors in the location of individual within a parish by using the parish centroid as the location of individual instead of using their address. In Table D.2, we see that the baseline results remain and the effects are similar in magnitude. We also control for the individual probability of being linked across censuses using a series expansion in Table D.3. The probability of being linked is based on the proportion of linked individuals within county-of-birth, census-year and name-frequency.

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 the relate contribution of each observations to our IV estimate, we compute the causal response weighting function following the decomposition proposed by Løken, Mogstad and Wiswall (2012). Figure D.4 shows weights along the proximity to the nearest train station. Figures D.5 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.6), rural/urban divide (Table D.5), age of fathers and sons (Tables D.7 and D.8), natives/foreigners (Table D.9), locals/outsiders (Table D.10) or farming occupation (Table D.11).

6 Mechanisms

The previous section presented causal empirical evidence that infrastructure in the form of railroad network led to an increase in intergenerational occupation mobility or broke the link between father-son occupations. This section further investigate the channels at work. Did better connectivity lead to the spatial mobility of workers? Or did it improve local labour market prospects?

Access to the railroad network could have improved the economic opportunity of individuals by connected residents to better job opportunities further away, attracting better options locally and changed the relative benefit of moving. We can therefore decompose the effect access to the railroad network on intergenerational mobility between individuals who move away and those who stay locally:

$$\Pr(IM|train) = \Pr(IM|stay, train) \times \Pr(stay|train) + \Pr(IM|move, train) \times \Pr(move|train) \tag{5}$$

where IM stands for intergenerational mobility and $train$ refers to the access to the railroad network (in our setting, this is measured as being one standard deviation closer to the nearest train station). The variables $move$ and $stay$ represent the individuals who have moved away from the county where they grew up and those who have stayed respectively. Taking the total derivative with respect to train, we obtain:

$$\begin{aligned}\Delta \Pr(IM|train) = & \underbrace{\Delta \Pr(IM|stay, train)}_{\text{change in IM induced by the train}} \\ &+ \underbrace{[\Delta \Pr(IM|move, train) - \Delta \Pr(IM|stay, train)]}_{\text{change in the returns to spatial mobility induced by the train}} \times \underbrace{\Pr(move|train)}_{\text{baseline spatial mobility}} \\ &+ \underbrace{[\Pr(IM|move, train) - \Pr(IM|stay, train)]}_{\text{baseline returns to spatial mobility}} \times \underbrace{\Delta \Pr(move|train)}_{\text{change in the spatial mobility from the train}}\end{aligned} \tag{6}$$

The train therefore affect the change in intergenerational mobility through three channels: (1) local opportunities, (2) the change in the returns to spatial mobility, and (3) the ease in spatial mobility. On the one hand, the railroad network affected the local economic activity. New industries with new job opportunities demanding new skills were established, this may have decoupled the ties between parents and their children's outcomes. On the other hand, the railroad network could have also affected upward mobility through spatial mobility. Railroads facilitated migration not only because they dramatically reduced travel time and cost but also because they likely increased information flows across connected districts. Sons moved away from the place where they grew up to find better opportunity elsewhere. Finally, moving would only have taken place if the relative benefits of moving outweigh the benefits of staying. The railroad would have changed the relative benefit of moving given the changes in opportunities induced by the railroad locally and further away. In the following section, we estimate each component of equation 6 to decompose the relative importance of each channel.

6.1 Returns to spatial mobility

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. For instance, movers may have earned more than those who stayed in their childhood parish because the brightest would have earned more regardless of their location were also most likely to move.

Following Abramitzky et al. (2012), we focus on sons who grew up in the same household. By comparing the outcome of sons who decided to move to their brothers who stayed, the estimate eliminates the across-family component of geographic mobility selection. This captures different propensities to move due to family background that are correlated both with the probability of spatial mobility and intergenerational mobility. Family background characteristics include financially constraints and poor local economic opportunities. It also eliminates the component of unobserved individual ability that is shared between brothers. We therefore estimate the following equation

$$f(Rank^{son}_{i,f,t}, Rank^{father}_{i,f,t}) = \tau_1 Mover_{i,f,t} + \tau_2 Proximity_{i,f,t-1} \times Mover_{i,f,t} + \phi_f + \epsilon_{i,f,t} \quad (7)$$

where i, f, and t index individual son, family and census year respectively. The dependent variable takes the same four measures as previously. The variable $Mover_{i,f,t}$ is an indicator variable equal to one if the son move away from the county where he grew up and $Promixity_{i,f,t-1}$ is the proximity to the nearest train station. The family fixed effect ϕ_f takes into account all within-family characteristics mentioned above. The coefficient τ_1 represents the change in baseline returns to spatial mobility. The coefficient τ_2 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 $Mover_{i,f,t}$.

In Table 7 we see that there is a significant and positive return to geographic mobility for all measures of intergenerational mobility. In other words, brothers who moved are less tied to their father's occupation although this can lead to both upward and downward mobility. However, the change in the returns to geographic mobility induced by better access to the railroad network is negative. That is, for the brothers who moved being better connected to the railroad network decreases intergenerational mobility. While they were more likely to follow their father's occupation and stay in the same occupational rank, they were also less downward mobile.

6.2 Spatial mobility pattern

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 have been shown to facilitate the spatial mobility for connected individuals by reducing the travel time and cost. By the time the South Eastern Railway opened as far as Dover, in 1844, 2210 miles of line had been opened, making travel around the country faster, more comfortable and less expensive. 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

$$\Pr(Mover_{i,c,t}) = \phi_1 Proximity_{i,c,t-1} + \phi_2 X_{i,c,t-1} + \gamma_t + \rho_c + \epsilon_{i,c,t} \quad (8)$$

where $mover_{i,c,t}$ 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.

Table 8 shows that railroads enabled individuals to move physically. There was a significant increase in the probability of moving away by 9 percentage points for sons who grew

Table 7: Returns to spatial mobility

	(1)	(2)	(3)	(4)
	$HISCO^{son} \neq HISCO^{father}$		$\lvert HISCAM^{son} - HISCAM^{father}\rvert$	
$\text{Mover}_{i,c,t-1}$	0.103***	0.068***	1.459***	1.031***
	(0.003)	(0.005)	(0.051)	(0.091)
$\text{Mover}_{i,c,t-1} \times \text{Proximity}_{i,c,t-1}$		−0.060***		−0.722***
		(0.006)		(0.129)
	(5)	(6)	(7)	(8)
	Upward Mobility		Downward Mobility	
$\text{Mover}_{i,c,t-1}$	0.040***	0.038***	0.016***	0.002
	(0.002)	(0.004)	(0.002)	(0.004)
$\text{Mover}_{i,c,t-1} \times \text{Proximity}_{i,c,t-1}$		−0.003		−0.024***
		(0.005)		(0.005)
Obs.	337,882	337,882	337,882	337,882

Notes: 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 HISCAM occupational 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 4 and 6/7 and 8). $Mover_{i,f,t}$ is an indicator variable equal to one if the son move away from the county where he grew up and $Promixity_{i,f,t-1}$ 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 7, all regressions include family fixed effects. The instrument consists of the interaction between our DLCP instrument and geographic mobility. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

up 5km closer to the train station. 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. Migration captures any costs related to being away from friends and family (e.g. return visits which are costly in terms of time and money) as well as any costs of not being able to consume the same types of goods as at home.[26] Bogart, Xuesheng, Alvarez, Satchell and Shaw-Taylor (2020) find similar results at the aggregate level. They 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.

[26]Geographical 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. By 1864, the scope of Law of Settlement had been greatly attenuated (Feldman, 2003). The nineteenth century has been characterised by a high rate of internal mobility, particularly from rural to urban areas.

Table 8: Geographic Mobility

	(1)	(2)
	$\Pr(Mover_{i,c,t})$	
Sample	All	Brothers
$\text{Proximity}_{i,c,t-1}$	0.091***	0.106***
	(0.021)	(0.025)
F-Stat	16.060	14.295
Avg. dep. var.	0.32	0.30
Obs.	969,242	337,882

Notes: The coefficients represent standardised $\text{Proximity}_{i,c,t-1}$ 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_{t-1} fixed effects 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 and port, the number of servants, household size and whether the father is born outside England and Wales. Column 1 includes our baseline sample of sons who are 40-52 years old and their father is observed 30 years earlier. Column 2 restricts the sample to brothers. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

6.3 Decomposition

In Table 9 we decompose 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 labour market opportunities. In particular, local opportunities account for roughly 90% of the upward mobility from being better connected to the railroad network, while the change in geographic mobility accounts for 8% and the change in the relative benefit from moving accounts for 2%.[27] Local opportunities include new job opportunities brought in by the train but also jobs opportunities becoming "local". The railroad offered the possibility

[27]This remains a decomposition exercise. Although we address the endogeneity issue in the decision to move, we do not take into account the destination location. The destination location is likely correlated with the individual's skill set and the complementarities in the labour 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. This may generate positive or negative spillovers to other areas.

of commuting thereby creating a separation between the place of work and place of residence thereby allowing workers to live away from their place of work, and consequently enlarging their employment possibilities (Heblich, Redding and Sturm, 2020). The improvement of local opportunities brought by better connectivity is consistent with well-known agglomeration effects, in which the dense population of urban areas has an effect on the productivity of resources. Alvarez et al. (2017) finds that the extension of the railroad in nineteenth century England and Wales led to population and employment growth.[28]

Table 9: Decomposition

	(1) $\text{HISCO}^{son} \neq \text{HISCO}^{father}$	(2) $\lvert HISCAM^{son} - HISCAM^{father}\rvert$	(3) Upward Mobility	(4) Downward Mobility
Proximity	0.072	1.029	0.047	0.005
	[0.049, 0.095]	[0.701, 1.350]	[0.035, 0.062]	[-0.008, 0.017]
Local opportunities	0.082	1.134	0.044	0.012
	[0.060, 0.104]	[0.816, 1.444]	[0.032, 0.060]	[-0.001, 0.025]
Ease of spatial mobility	0.007	0.109	0.004	0.000
	[0.004, 0.010]	[0.067, 0.151]	[0.003, 0.006]	[0.000, 0.001]
Returns to spatial mobility	-0.018	-0.214	-0.001	-0.007
	[-0.021, -0.015]	[-0.272, -0.160]	[-0.003, 0.001]	[-0.009, -0.005]

Notes: The sample consists of brothers who are 40-52 years old and their father observed 30 years earlier. We compute the standard errors by re-sampling by parish of origin to create a parish cluster bootstrap. Confidence intervals are based on 500 replications.
"Proximity" $= \Delta \Pr(IM|train) = \text{Total } (\hat{\alpha}_1)$ (see Table C.5)
"Local opportunities" $= \Delta \Pr(IM|stay, train) = \hat{\alpha}_1 - \hat{\beta}_1\hat{\phi}_1 - \hat{\beta}_2 \frac{\sum_t \sum_{i=N}^{N_t} mover_{i,f,t}}{\sum_t N_t}$
"Ease of spatial mobility" $= [\Pr(IM|move, train) - \Pr(IM|stay, train)] \times \Delta \Pr(move|train) = \hat{\beta}_1\hat{\phi}_1$
"Returns to spatial mobility" $= [\Delta \Pr(IM|move, train) - \Delta \Pr(IM|stay, train)] \times \Pr(move|train) = \hat{\beta}_2 \frac{\sum_t \sum_{i=N}^{N_t} mover_{i,f,t}}{\sum_t N_t}$

7 Conclusion

The long-run implications of infrastructure improvements are of interest both for historical reasons and also because they are related to current debates on institutional change. Many countries are currently investing or considering investing in large transport infrastructure improvements. Systematic evaluation of the long-term causal effects of large infrastructure projects can inform these important policy decisions. 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 a new dataset which allows us to observe the occupation of father-son pair between 1851 and

[28]We reproduce the findings from Alvarez et al. (2017) in Table C.7 in the Appendix. We observe that better connected parishes experience an increase in population density along with an increase in the average and median occupational score.

1911 and geographically locate them down to the street level. This new level of disaggregation allows us to measure access the railroad network using 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. Sons who grew up one standard deviation (approximately 5km or one hour's walk) closer to the nearest train station were 6 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 ove out of farming activities, but also transitions into industrial and commercial activities. This also resulted in an important shift in the distribution of occupational ranking. Better connected sons were more likely to move to either end of the occupation ranking. This significantly benefitted sons from the upper and lower classes, while sons in the middle class benefitted the least.

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 labour opportunities brought to town by the railroad or becoming feasible by commuting. This implies that when evaluating the effectiveness of transport infrastructure, focusing on those who move away will provide an underestimate. Our results also motivate place-focused approaches to improving economic mobility such as making investment to improve outcomes in areas that currently have low levels of mobility or providing access to affordable transportation.

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.

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.

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

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.

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

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 .

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., 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.

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

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.

Duflo, E., Pande, R., 2007. Dams. The Quarterly Journal of Economics 122, 601–646.

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.

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

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

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.

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

Jantti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Bjorklund, A., Eriksson, T., 2006. American exceptionalism in a new light: A comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States .

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.

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.

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.

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 .

Schürer, K., Higgs, E., 2014. Integrated census microdata (I-CeM); 1851-1911 [computer file]. Colchester, Essex: UK Data Archive [distributor], April 2014. SN: 7481.

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.

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.

Appendices

A Data Construction

A.1 Data Sources

I-CEM The I-CeM project, lead by Professor Kevin Schurer and Professor Eddy Higgs, digitalized and standardised, and coded the England and Wales census of 1851, 1861, 1881, 1891, 1901 and 1911. The full name and address can be accessed via special license.

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

Parish and county boundaries provided by the UK Data Service.

HISCAM 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.[29] 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.

Railways of Great Britain GIS shapefiles of railways lines and stations from 1851 and 1881 from England, Wales and Scotland, digitised by the Cambridge Group for the History of Population and Social Structure. This was digitised 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)

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

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

Figure A.1: Roman Roads

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).

City types This is taken from the British Parliamentary Papers (HC348, 1831) (Casson, 2009) Table 3.3.

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, standardised 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 based 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.

Table A.1: Linkage Statistics

	County	Std. parish	Parish
		1851-1881	
Step 1	4,164,488	2,158,059	1,850,017
Step 2	828,946	1,427,241	1,208,746
Step 3	640,319	214,777	171,155
Step 4	1,208,917	1,571,511	1,329,712
Linkage rate	15	19	16
		1881-1911	
Step 1	6,996,906	3,961,464	2,781,673
Step 2	1,537,250	2,626,026	1,912,978
Step 3	1,099,825	429,448	269,452
Step 4	2,147,941	2,905,267	2,094,985
Linkage rate	17	23	17

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).

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.

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

	1851-1881	1881-1911
Nb. individuals	652,192	1,227,324
Linkage rate	42	49
Avg. age distance	0.54	0.41
Avg. surname Jaro-Winkler distance	0.01	0.01
Avg. name Jaro-Winkler 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).

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 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 40-52 years old. This implies that when the father's occupation was measured, 30 years

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

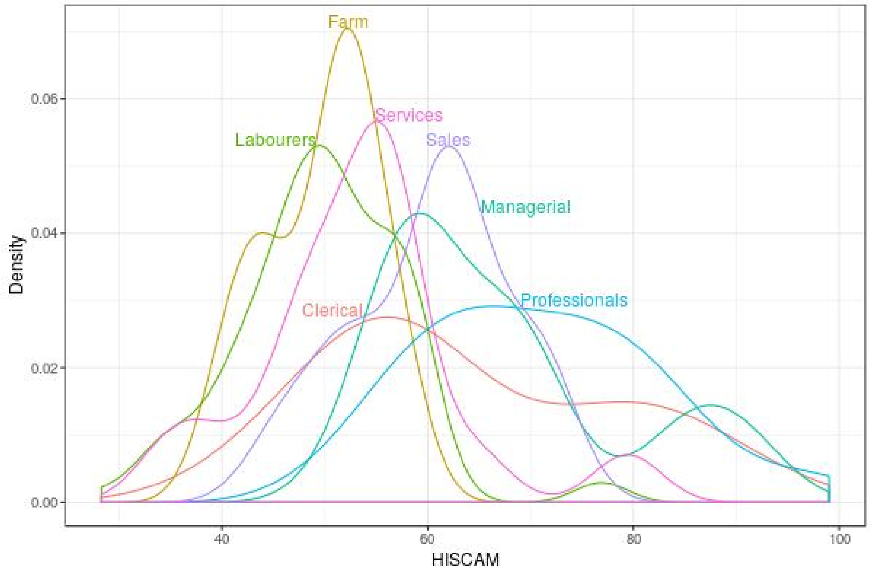
Article	Source	Match rate	Number linked
Costas Fernandez et al. (2020)	1881 England and Wales Census to 1911 England and Wales Census (Full, Men 40-52)	49%	1,227,324
Costas Fernandez et al. (2020)	1851 England and Wales Census to 1881 England and Wales Census (Full, Men 40-52)	42%	652,192
Guerra and 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)

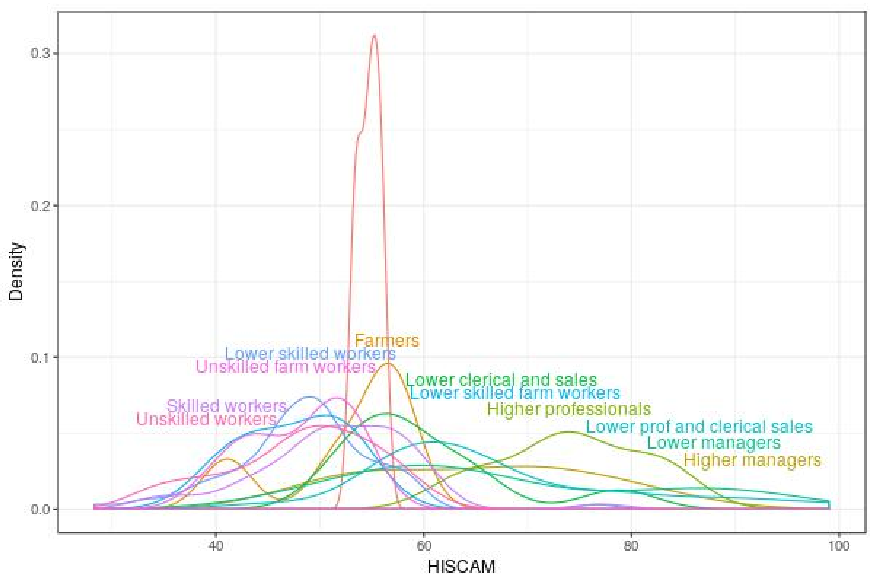
earlier, the son was 10-22. Moreover, if in any of these father-son pairs has a Jaro-Winkler distance between father and son surname larger than 0.12 we disregard it.

A.4 Occupation classification

Figure A.2: Density of HISCAM occupation ranking by HISCO and HISCLASS occupation classification



(a) HISCO



(b) HISCLASS

A.5 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).[30] 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.[31]

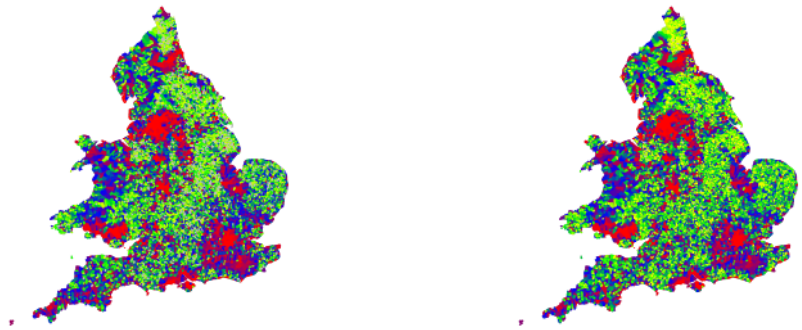
Whenever we use information at the parish level for 1911 we need to standardize the parish definition. This is because the I-CEM data provides a parish division of the UK that is homogeneous for the 1851 and 1881 censuses. However, in the 1911 this division changes. For example, Central London in the 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`.)

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

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

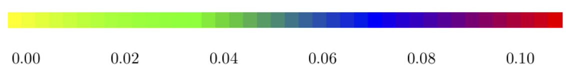
A.6 Descriptives of linked sample

Figure A.3: Size of the sample and population



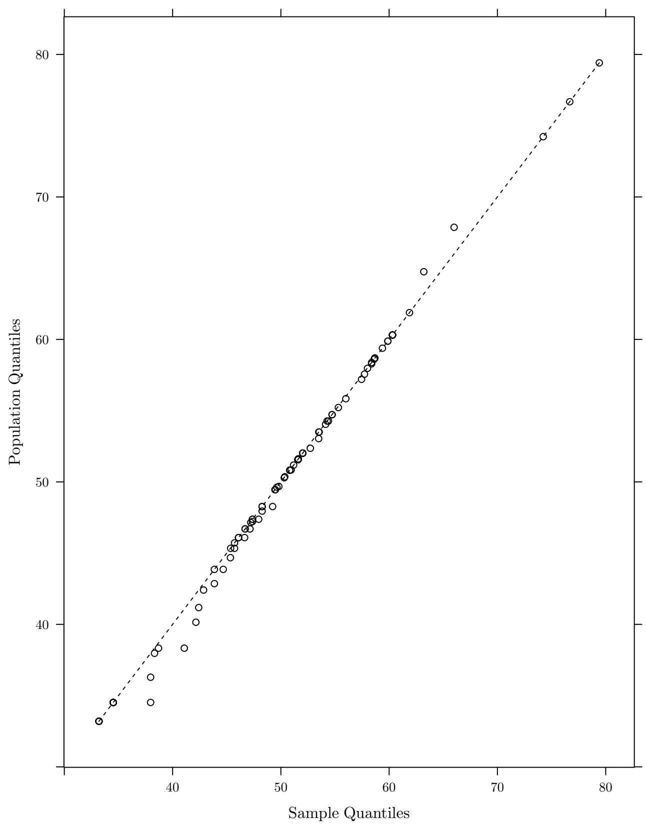
(a) Sample

(b) Census



Note: Figure A.3a displays the sample sizes in our main dataset, i.e. linked males that are 40-52 year old at the parish where they currently live. Sample sizes computed by pooling years 1881 and 1911 for every parish. Figure A.3b displays the parish populations of males aged 40-52 pooling data from 1881-1911. Sizes are represented as percentage of the total. The legend covers the 1 to 99 percentile. Parishes that could not be uniquely matched across censuses are in grey.

Figure A.4: HISCAM Distribution
Census vs Matched Sample



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 with a valid occupation code that is matched to HISCAM occupational rank. Both distributions are constructed by pooling the 1881 and 1911 censuses.

Table A.4: Descriptive Statistics for the linked and non-linked samples

	Non-Linked	Linked	T-Statistic (Difference)
Avg. age	45.855	44.702	288.359
Name freq.	0.054	0.053	20.379
Surname freq.	0.001	0.001	92.804
Share of foreign Born	0.122	0.029	365.247
Avg. HISCAM occ. rank	49.962	50.049	-7.690
Share of professional	0.041	0.030	52.890
Share of managerial	0.022	0.021	8.172
Share of clerical	0.042	0.044	-8.802
Share of sales	0.105	0.102	9.148
Share of services	0.063	0.053	41.332
Share of agricultural	0.136	0.160	-61.707
Share of labourers	0.591	0.590	2.078
Obs.	2,681,281	1,183,071	

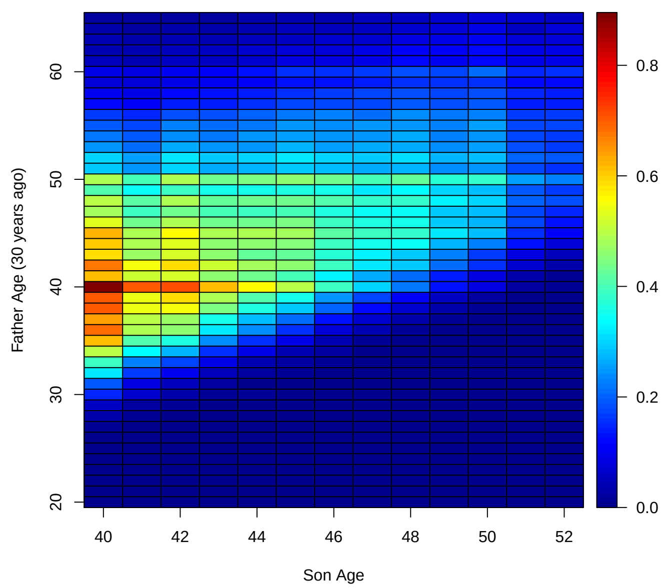
Note: The "non-linked" sample includes all men aged 40-52 that have not been linked (not in our sample). The "linked" sample includes our estimation sample (i.e. men aged 40-52).

Table A.5: Role of railroad network access on linked sample

	Dep. var.: Share of linked individuals among the parish population aged 10-22	
	(1)	(2)
	DLCP network	Nearest train station
Proximity$_{p,c,t}$	-0.001	-0.002^{**}
	(0.001)	(0.001)
Obs.	24,450	24,450

Notes: Each coefficient represents the coefficient of the standardised Proximity$_{p,c,t-1}$ 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 in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

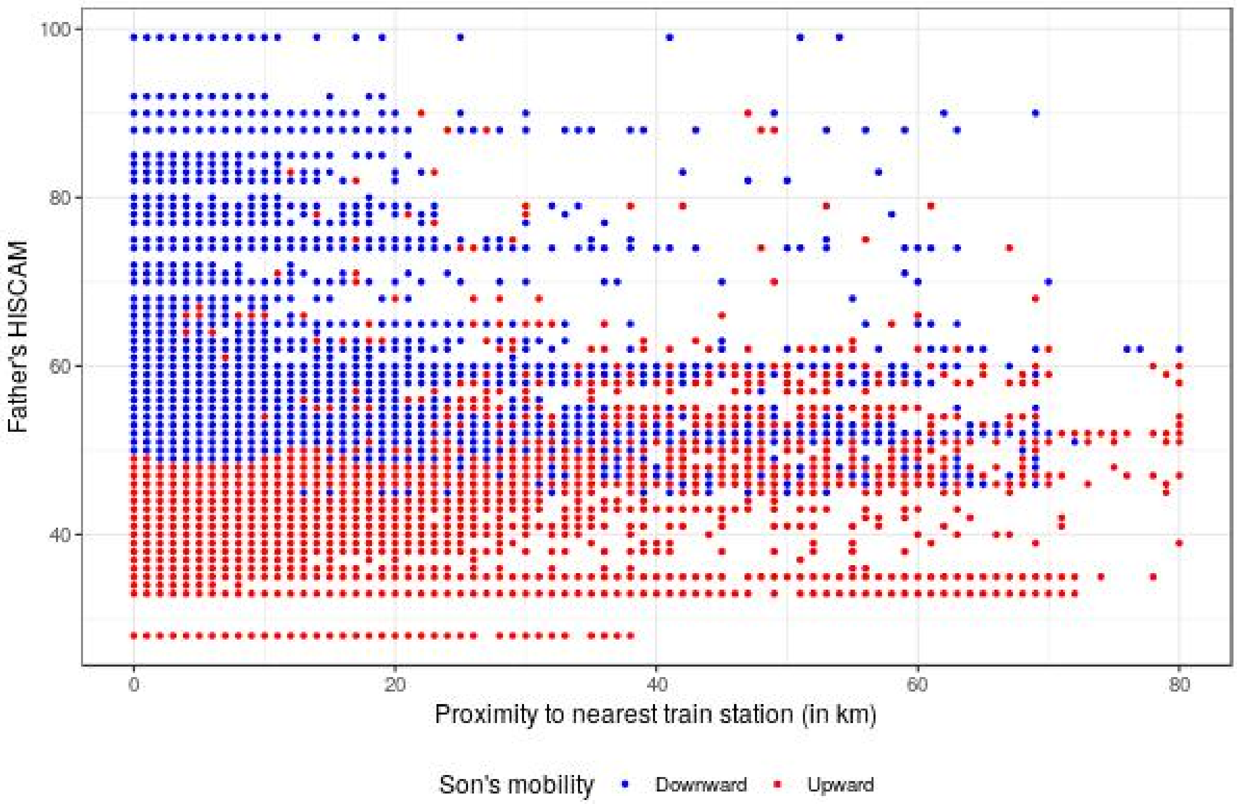
Figure A.5: Joint distribution of father-son ages



Note: This figure depicts the joint distribution of the ages of fathers and sons in our linked sample.

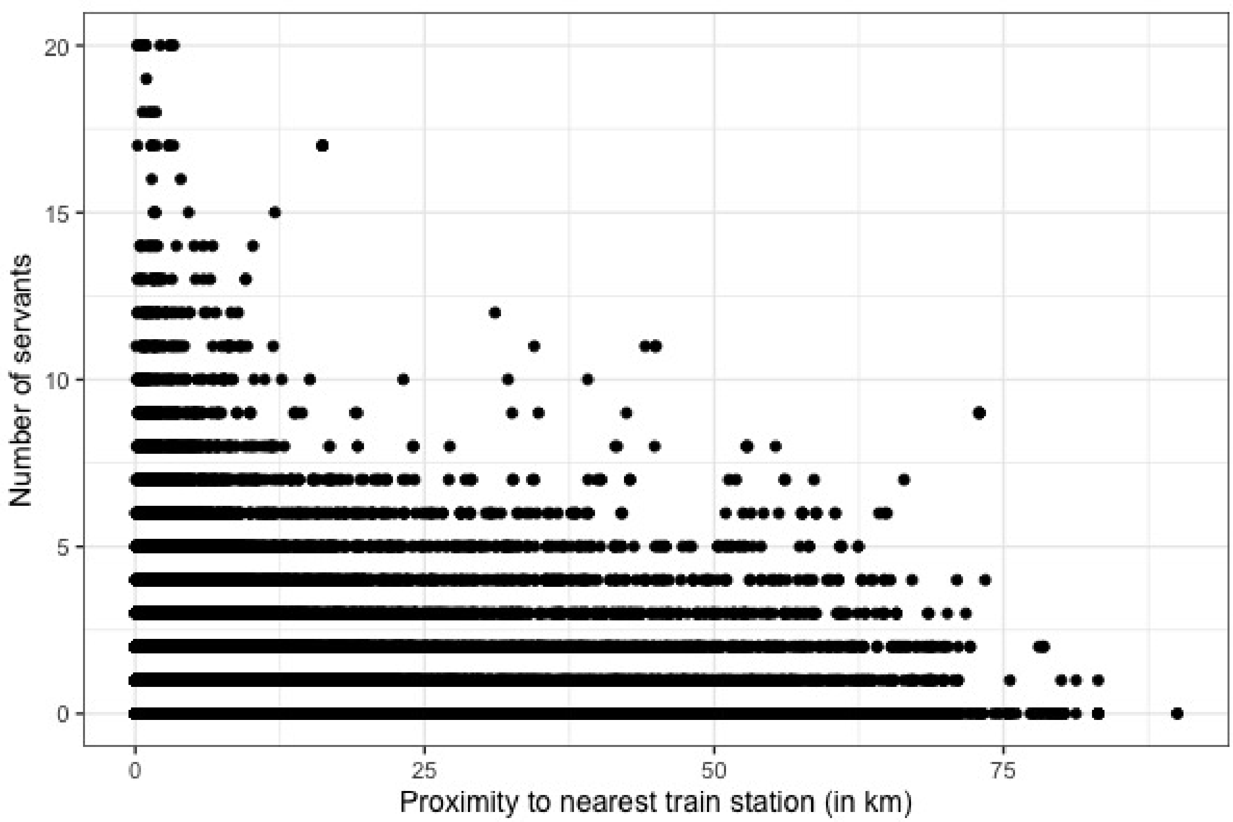
B Additional Descriptives

Figure B.1: Proximity to train station and intergenerational mobility



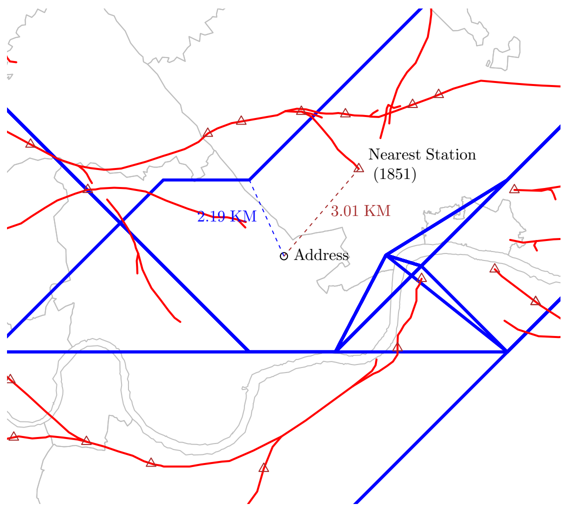
Note: This plot displays the relationship between the distance to the nearest train station during childhood and HISCAM of fathers. Colours represent the intergenerational mobility patterns of sons (red if there is a higher share of sons who are upward mobile than downward mobile, and blue otherwise).

Figure B.2: Proximity to train station and number of servants



Note: This plot displays the relationship between the distance to the nearest train station and the number of servants during the childhood.

Figure B.3: Example



Note: Red lines are the actual railroad lines while blue lines are the projected railroad lines. Triangles are train stations. An individual residing on the black dot is 3.01km from the nearest train station and 2.19km from the nearest projected railroad line.

Table B.1: Descriptives for Brother Sample

	Mean	St. Dev.	Min.	Median	P75	Max.
			A. BROTHERS			
Age	44.71	3.52	40	44	47	52
Foreign-born	0.03	0.16	0	0	0	1
Urban resident	0.39	0.49	0	0	1	1
Literate	0.29	0.45	0.00	0.00	1.00	1.00
HISCAM occupation rank	49.59	10.08	28.28	50.36	54.81	99.00
$HISCO^{son} \neq HISCO^{father}$	0.79	0.41	0	1	1	1
$\|HISCAM^{son} - HISCAM^{father}\|$	7.98	8.36	0	5.9	12.6	62
Upward mobility	0.16	0.37	0	0	0	1
Downward mobility	0.17	0.38	0	0	0	1
County mover	0.32	0.47	0	0	1	1
Dist. to nearest train station (in km)	3.65	6.08	0.01	1.60	4.08	83.17
Dist. moved | county mover	102.69	100.06	0.06	71.82	152.56	628.55
			B. FATHERS			
Age	47.05	7.57	20	47	52	65
Foreign-born	0.06	0.24	0	0	0	1
Urban resident	0.38	0.49	0	0	1	1
Household size	6.77	2.15	0	7	8	18
Number of sons	4.65	2.09	0	5	6	14
Number of servants	0.19	0.70	0	0	0	39
HISCAM occupation rank	49.64	8.96	28.28	51.18	53.50	99.00
Literate	0.30	0.46	0.00	0.00	1.00	1.00

Note: The sample consists of 77,407 sons from 35,297 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.2: Change in the share of occupations 1851-1911

HISCO		% in 1911	% in 1851
	Top 5 Declining occupations		
62110	Farm workers, specialisation unknown	3.77	18.45
61110	General farmers and farmers nfs	1.98	4.50
80100	Boot and shoe makers and repairers	1.39	3.57
75400	Weavers	0.86	2.39
79120	Tailors and tailoresses	0.75	1.90
	Top 5 growing occupations		
98550	Delivery men and drivers of goods	2.30	1.32
84130	Machine makers, builders and fitters	1.62	0.20
41010	Dealer, merchant etc. (Wholesale and retail trade)	6.71	4.77
39310	Office clerks, specialisation unknown	3.35	0.78
71120	Miners	7.46	4.25

C Additional Results

Table C.1: The Effect of Rail Connection by Woollard Occupations Classification

	(1)	(2)	(3)	(4)	(5)	(6)
	Father					
	Professional	Industrial	Commercial	Domestic	Agriculture	All
Professional	-0.027	0.011^{**}	0.007	0.015	-0.001	0.012^{***}
	(0.028)	(0.004)	(0.013)	(0.022)	(0.004)	(0.005)
Industrial	0.044	0.011	0.044	0.073^{*}	0.092^{***}	0.099^{***}
	(0.031)	(0.011)	(0.032)	(0.040)	(0.015)	(0.014)
Commercial	0.025	0.023^{***}	0.047	0.057^{*}	0.013	0.025^{***}
	(0.024)	(0.007)	(0.029)	(0.030)	(0.008)	(0.007)
Domestic	-0.001	-0.006^{**}	-0.018^{*}	-0.040	0.020^{***}	-0.003
	(0.009)	(0.003)	(0.010)	(0.030)	(0.005)	(0.003)
Agriculture	-0.048^{***}	-0.047^{***}	-0.088^{***}	-0.111^{***}	-0.125^{***}	-0.139^{***}
	(0.015)	(0.007)	(0.023)	(0.030)	(0.018)	(0.014)
Obs.	30,022	613,244	90,235	16,270	216,839	969,242

Notes: Each entry represents the coefficient of the standardised Proximity$_{i,c,t-1}$ instrumented by the proximity to the DLCP network. The dependent variable is an indicator equal to one if the son works in a specific Woolward occupation (rows). Observations include sons who are 40-52 years old and their father's occupation is measured 30 years earlier (column 6). Additional sample restriction are that the fathers work as "professional" (column 1), "industrial" (column 2), "domestic" (column 3), "commercial" (column 4), and "agriculture" (column 5). All regression include census year and childhood county$_{t-1}$ 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 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. *p<0.1; **p<0.05; ***p<0.01

Table C.2: Skill level based on occupation

	(1) Father illiterate	(2) Father unskilled	(3) All
Son literate	0.070***	0.096***	0.081***
	(0.010)	(0.016)	(0.015)
Son skilled	0.050***	0.041***	0.042***
	(0.013)	(0.011)	(0.013)
Obs.	661,406	705,937	966,732

Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$ instrumented by the proximity to the DLCP network. 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 (column 3). The sample includes sons whose fathers are illiterate (column 1) and unskilled (column 2). Given that some occupations are not linked to any literacy requirement, we lose some observations with respect to the baseline sample. All regressions include county and census 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 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. *p<0.1; **p<0.05; ***p<0.01

Table C.3: The effect of rail connection by HISCLASS occupations

	(1)	(2)	(3)	(4)	(5)
			Father		
	Manager	Skilled Worker	Lower Skilled	Farmer	Any
Manager	0.013	0.002	0.003	0.008	0.014*
	(0.022)	(0.007)	(0.008)	(0.007)	(0.007)
Skilled Workers	0.030*	-0.026^{**}	0.028**	0.028***	0.042***
	(0.015)	(0.013)	(0.014)	(0.008)	(0.010)
Lower Skilled	0.039*	0.062***	0.046**	0.075***	0.081***
	(0.017)	(0.014)	(0.019)	(0.012)	(0.013)
Farmers	-0.082^{***}	-0.039^{***}	-0.077^{***}	-0.111^{***}	-0.137^{***}
	(0.018)	(0.007)	(0.012)	(0.017)	(0.015)
Obs.	103,456	222,899	356,653	220,520	928,121

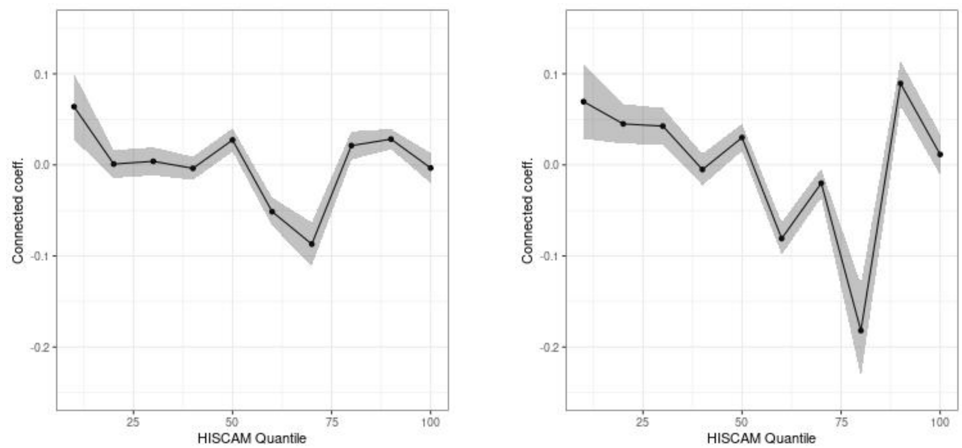
Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$ to the nearest train station instrumented by the proximity to the DLCP network. The dependent variable is an indicator equal to one if the son works in a specific HISCLASS occupation (rows). Observations include 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-1} 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 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. *p<0.1; **p<0.05; ***

Table C.4: White vs blue collar occupations

	(1) Father in white collar occ.	(2) Father in blue collar occ.
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.030^{***}	0.058^{***}
	(0.011)	(0.011)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	-0.234	1.068^{***}
	(0.372)	(0.138)
Upward Mobility	0.015	0.055^{***}
	(0.010)	(0.007)
Downward Mobility	-0.019	0.001
	(0.015)	(0.004)
Obs.	168,455	800,787

Notes: Each coefficient represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$ 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 HISCAM 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. Sample is further restricted based on the type of occupation held by the father: white collar (column 1; HISCO 0 to 5) and blue collar (column 2; HISCO 6 to 9). All regressions include fixed effects for census year and childhood county_{t-1}. 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. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

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



(a) Sons (b) Fathers

Note: Each dot represent the coefficient of the standardised Proximity$_{i,c,t-1}$ 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 childhood county$_{t-1}$. 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.5: Social Mobility Pattern by Brothers

	(1)	(2)	(3)
$HISCO^{son} \neq HISCO^{father}$	0.102***	0.073***	0.072***
	(0.012)	(0.015)	(0.014)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	1.296***	1.053***	1.029***
	(0.175)	(0.215)	(0.213)
Upward Mobility	0.053***	0.048***	0.047***
	(0.007)	(0.009)	(0.009)
Downward Mobility	0.009	0.006	0.005
	(0.006)	(0.008)	(0.008)
SW-F	118.731	21.397	12.865
F-Stat	118.731	128.383	128.651
Obs.	337,882	337,882	337,882
Year FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel routes	No	Yes	Yes
Household characteristics	No	No	Yes

Notes: Each cell represents the coefficient of the standardised Proximity$_{i,c,t-1}$ 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 HISCAM 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 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$_{t-1}$. 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 and port (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). F-Stat reports Sanderson and Windmeijer (2015) weak instrument F-statistic. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table C.6: Growing/declining industries

	(1)	(2)	(3)
	Occupation of father		
	Growing	Declining	Any
Growing	0.066^{*}	0.044^{***}	0.050^{***}
	(0.035)	(0.006)	(0.006)
Declining	-0.146^{***}	-0.102^{***}	-0.147^{***}
	(0.032)	(0.012)	(0.015)
SW-F	6.501	16.716	14.454
F-Stat	65.009	167.165	144.540
Obs.	75,291	454,079	969,242

Notes: Growing/declining is an indicator variable is an individual works in a HISCO occupation within the top/bottom 25% of the growth industry (see Table B.2 for examples). The growth of industry is based on the difference in the share of individuals in a HISCO occupation between 1851 and 1911. All regressions include fixed effects for census year and childhood county$_{t-1}$. 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. F-Stat reports Sanderson and Windmeijer (2015) weak instrument F-statistic. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table C.7: Parish level results

	(1)	(2)
Proximity$_{p,1851}$		
Δ log(Population density)$_{p,1881}$	0.999***	0.862***
	(0.185)	(0.203)
HISAM$^{mean}_{p,1881}$	1.552*	1.939**
	(0.836)	(0.989)
HISCAM$^{median}_{p,1881}$	0.162	0.569
	(1.119)	(1.309)
HISCAM Gini$_{p,1881}$	0.044***	0.031***
	(0.009)	(0.009)
County FE	Yes	Yes
Historical importance of town	No	Yes
Historical travel routes	No	Yes
SW-F	62.260	7.743
F-Stat	62.260	46.456
Obs.	11,125	11,125

Notes: Each cells represents the coefficient of the standardised Proximity$_{p,1851}$ between the centroid of the parish and the nearest train station in 1851, instrumented by the indicator whether the parish is connected to the DLCP railroad network. The dependent variable is the change in the log population density between 1851 and 1881 (row 1), the average HISCAM (row 2), the median HISCAM (row 3), and the HISCAM Gini (row 4). The sample includes the 11,125 parishes in 55 counties in 1881. All regressions include county fixed effects. 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, and the distance to the closest Roman road and port (column 2). Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

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 true railroad network. Figure D.1 shows that our baseline results are conservative.

Alternative measure of 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. 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 show the results once we add higher polynomials to the control variables and parish fixed effects. The parish fixed effect controls for very local characteristics such as local public goods. The proximity coefficient remains significant and of similar magnitude. Moreover, the coefficients between the baseline and these alternative specifications are not statistically significant. 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. When including parish fixed effect, the effect of proximity to the railroad network on occupational ranking becomes smaller in magnitude but still positive and significant. The effects becomes similar in magnitude when looking at the occupational categories.

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

$$f(Rank^{son}_{i,c,t}, Rank^{father}_{i,c,t-1}) = \sum_{d=1}^{\overline{d}} \beta_d \text{DProx}_{d,i,c,t-1} + \alpha_2 X_{i,c,t-1} + \gamma_t + \rho_c + \nu_{i,c,t-1}$$

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.4 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:

$$\begin{aligned} f(Rank^{son}_{i,c,t}, Rank^{father}_{i,c,t-1}) = {} & \theta_1 Proximity_{i,c,t-1} + \theta_2 (Proximity_{i,c,t-1})^2 \\ & + \theta_3 X_{i,c,t-1} + \gamma_t + \rho_c + \epsilon_{i,c,t-1} \end{aligned} \tag{9}$$

We use the square distance to the hypothetical railroad network as an instrument for ($\text{Proximity}^2_{i,c,t-1}$). Figures D.5 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.6 shows us 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 and rural area separately in Table D.5. We do not observe large differences between the two groups. Individuals living in an urban area is defined as those who grew up within 2.5km of a 1801 town.

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.

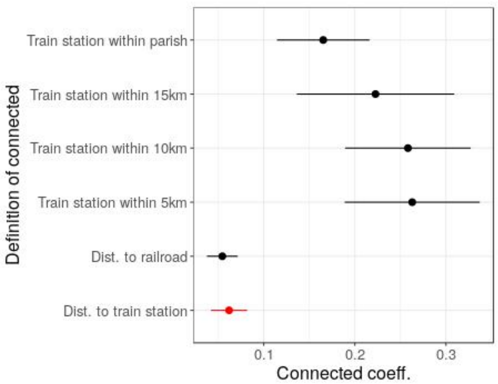
Age As several studies have shown (e.g. (Grawe, 2006)), estimates of intergenerational mobility is highly sensitive to the age at which sons' labour market outcomes are observed, increasing substantially in age. This can be explained by the strong life-cycle pattern in the correlation between current 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 at $t-1$ 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.

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.9 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.

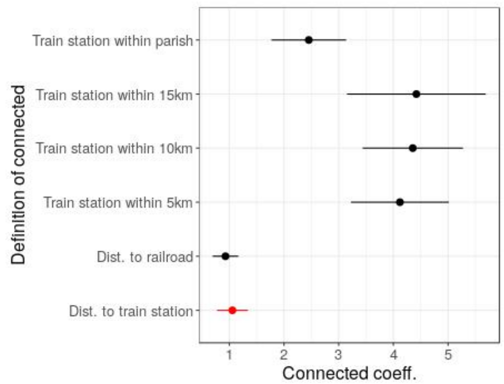
Locals vs Outsiders The estimator would also be biased if people and firms move over time along the same spatial lines as the forecasted placement of the railroad network. For instance, fathers who have high ambition for their family may decide to live in connected parishes. We explore the possibility of self-selection in two ways. We first estimate our regression for fathers who were born in the parish they are currently living in (i.e. stayers) and those who have moved (i.e. movers). In Table D.10, we see that intergenerational mobility patterns can be seen for both stayers and movers.

Farming activities Farming occupations may have a particular role in the transition during the Industrial Revolution. In Table D.11 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.

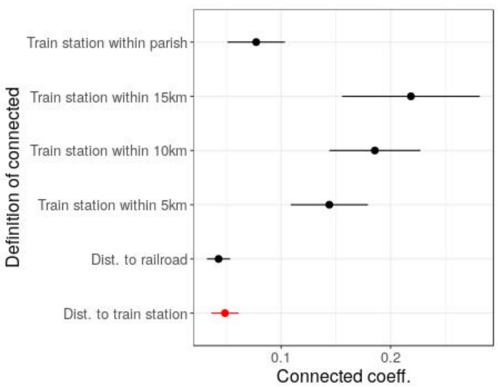
Figure D.1: Alternative definition of proximity



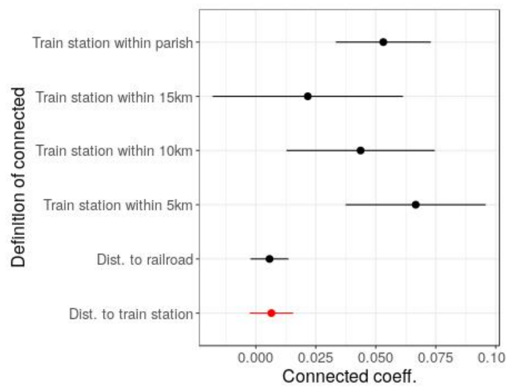
(a) $\text{HISCO}^{son} \neq \text{HISCO}^{father}$



(b) $|HISCAM^{son} - HISCAM^{father}|$



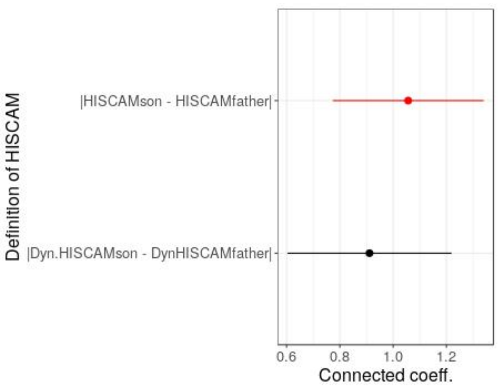
(c) Upward mobility



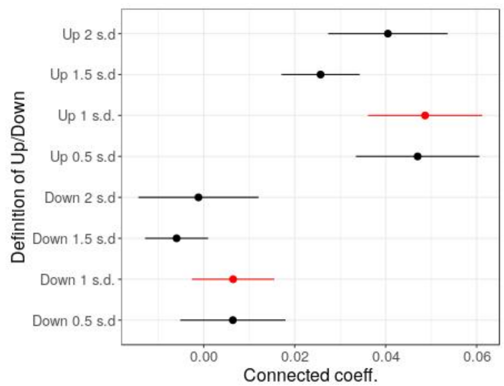
(d) Downward mobilty

Notes: Each dot represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, 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 fixed effects for census year and childhood county_{t-1} 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



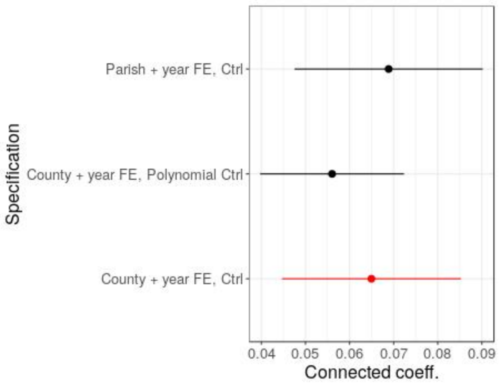
(a) Definition of HISCAM



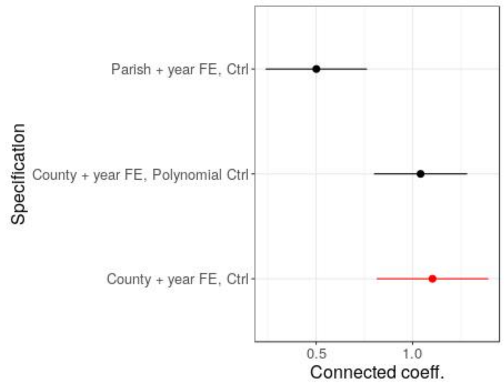
(b) Definition of Up/Down

Notes: Each dot represents the coefficient of the standardised Proximity$_{i,c,t-1}$, 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 fixed effects for census year and childhood county$_{t-1}$ 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

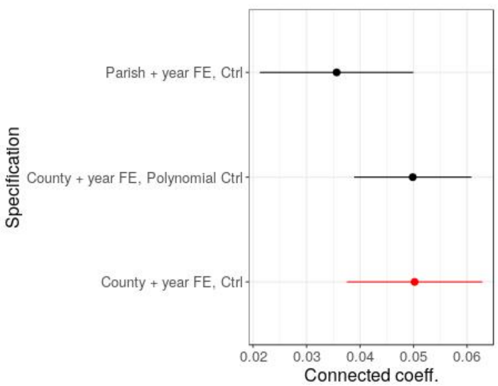
Figure D.3: Alternative specification



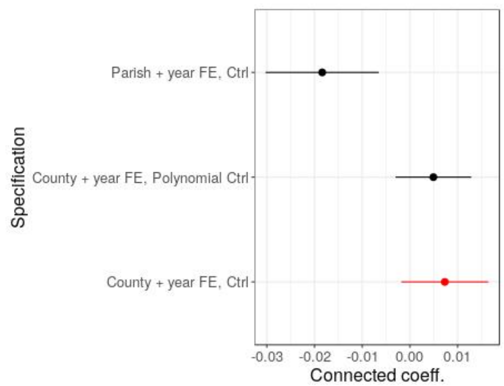
(a) $\text{HISCO}^{son} \neq \text{HISCO}^{father}$



(b) $|HISCAM^{son} - HISCAM^{father}|$



(c) Upward mobility



(d) Downward mobilty

Notes: Each dot represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, 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 childhood county_{t-1}, 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 in year $t-1$ 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)
$HISCO^{son} \neq HISCO^{father}$	0.091***	0.067***	0.064***
	(0.010)	(0.011)	(0.010)
$\lvert HISCAM^{son} - HISCAM^{father}\rvert$	1.259***	1.132***	1.085***
	(0.135)	(0.152)	(0.149)
Upward Mobility	0.051***	0.051***	0.049***
	(0.006)	(0.007)	(0.006)
Downward Mobility	0.013***	0.009*	0.008*
	(0.004)	(0.005)	(0.005)
Obs.	918,478	918,478	918,478
Year FE	Yes	Yes	Yes
County$_{t-1}$ FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel route	No	Yes	Yes
Household characteristics	No	No	Yes

Notes: Each entry represents the coefficient of the standardised Proximity$_{i,c,t-1}$, 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 HISCAM 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). Occupations ranking exclude all rail related occupations. 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$_{t-1}$. 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 and port (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 in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.2: Measurement error in geolocation

	(1)	(2)	(3)	(4)	(5)	(6)
		Address			Parish centroid	
HISCO^{son} $\neq$ HISCO^{father}	0.089***	0.065***	0.062***	0.091***	0.074***	0.071***
	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
$\lvert HISCAM^{son} - HISCAM^{father}\rvert$	1.249***	1.102***	1.057***	1.264***	1.138***	1.085***
	(0.132)	(0.147)	(0.144)	(0.136)	(0.138)	(0.136)
Upward Mobility	0.051***	0.050***	0.049***	0.052***	0.049***	0.047***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Downward Mobility	0.012***	0.007	0.006	0.012***	0.010**	0.009*
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Obs.		969,242			969,242	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County$_{t-1}$ 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

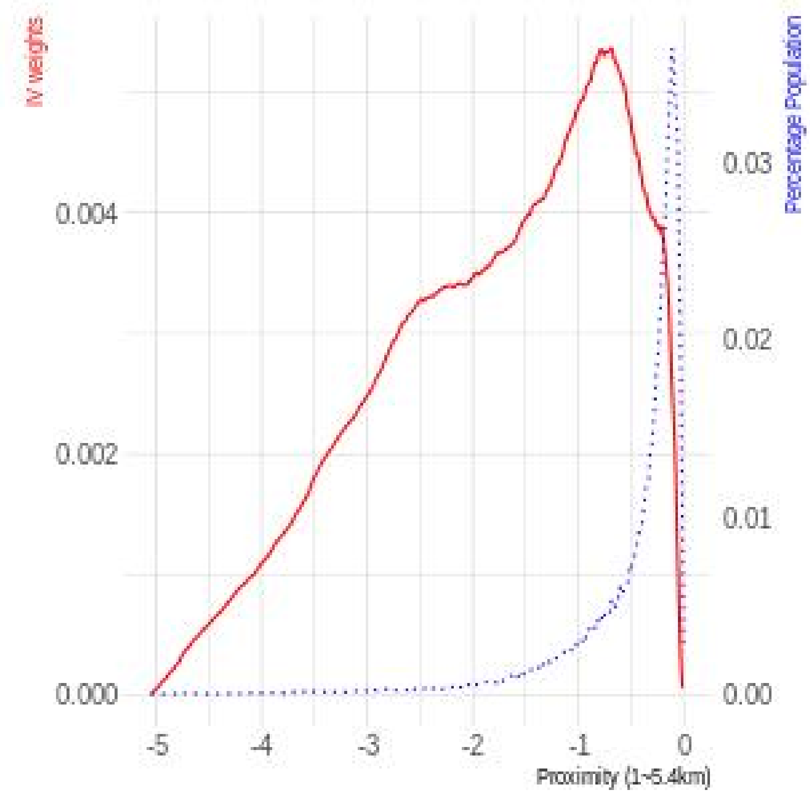
Notes: Each entry represents the coefficient of the standardised Proximity$_{i,c,t-1}$, instrumented by the proximity to the DLCP network. Individuals are geolocated based on their address (columns 1 to 3) or at the parish centroid (columns 4 to 6). The dependent variables are an indicator variable which 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 HISCAM 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 childhood county$_{t-1}$. 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 and port (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 in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.3: Controlling for the selection

	(1)	(2)	(3)
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.082***	0.057***	0.055***
	(0.009)	(0.010)	(0.010)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	1.137***	0.966***	0.987***
	(0.125)	(0.140)	(0.137)
Upward Mobility	0.048***	0.046***	0.046***
	(0.005)	(0.006)	(0.006)
Downward Mobility	0.010***	0.005	0.006
	(0.004)	(0.005)	(0.004)
SW-F	37.831	19.042	17.445
F-Stat	113.493	152.340	157.003
Obs.	969,242	969,242	969,242
Year FE	Yes	Yes	Yes
County_{t-1} FE	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes
Historical travel routes	No	Yes	Yes
Quadratic Prob.Linkage	Yes	Yes	No
Cubic Prob.Linkage	No	No	Yes

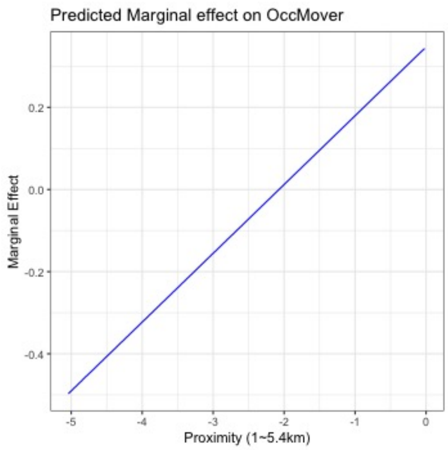
Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, instrumented by the proximity to the DLCP network. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute value of the difference in the HISCAM 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 childhood county_{t-1}. Additional controls include the quadratic probability of linkage (columns 1 and 2), 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 and port (columns 2 and 3), household characteristics consisting the number of servants, household size and whether the father is born outside England and Wales (column 3), and cubic probability linkage (column 3). Standard errors clustered at the parish in year $t-1$ are reported in parentheses. SW-F is the Sanderson and Windmeijer (2016) F-statistic for weak instruments. *p<0.1; **p<0.05; ***p<0.01.

Figure D.4: IV weights

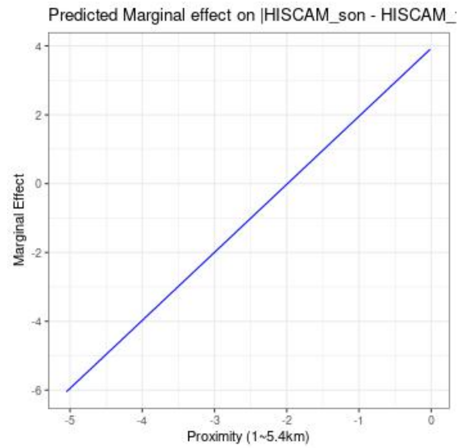


Note: This figure shows the population share (right axis) and the assigned weights in the IV estimates (left axis) over the proximity to the nearest train station. The x-axis represents units (5.4 km each) of proximity to the nearest train station winsorised at the 1%.

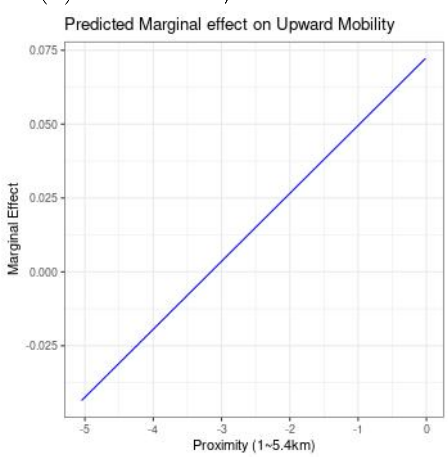
Figure D.5: Predicted marginal effect



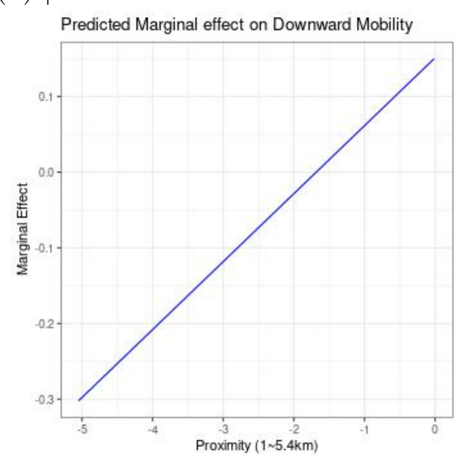
(a) HISCO^{son} $\neq$ HISCO^{father}



(b) $|HISCAM^{son} - HISCAM^{father}|$



(c) Upward mobility



(d) Downward mobility

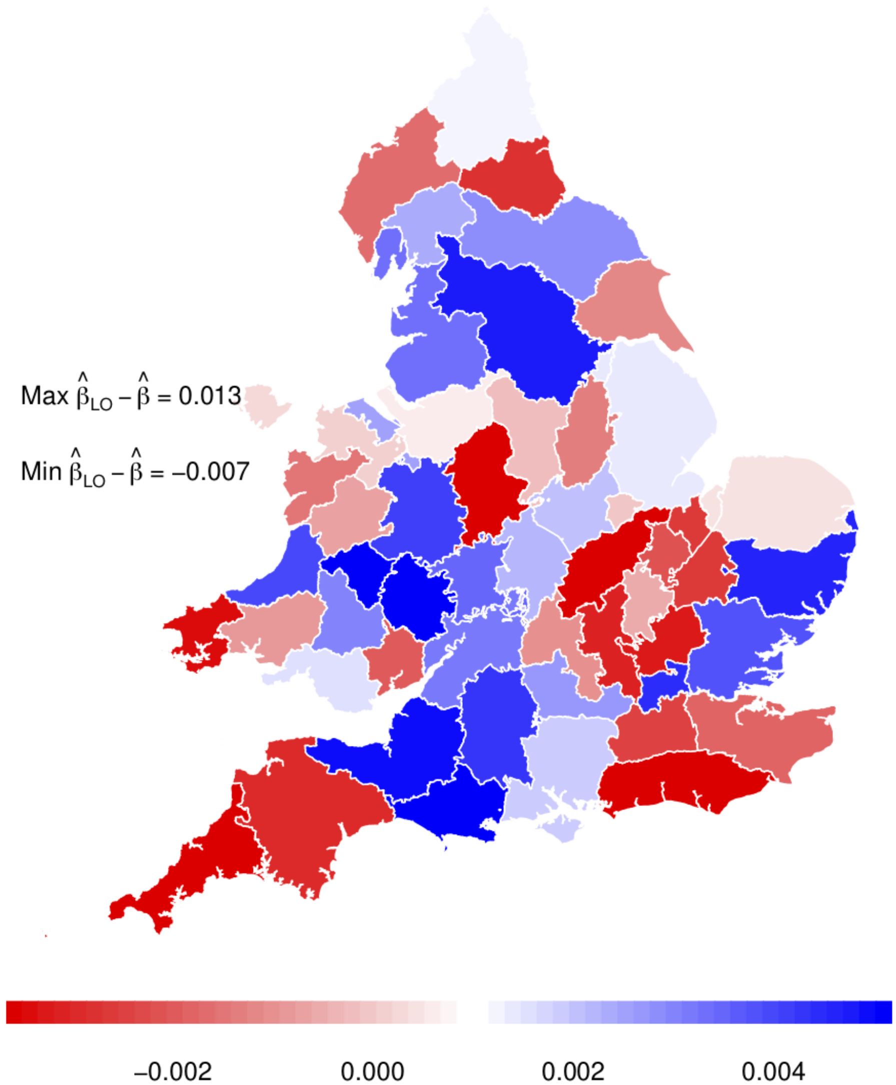
Note: This figure presents the predicted marginal effect of equation 9. The x-axis represents units (5.4 km each) of proximity to the nearest train station winsorised at the 1%.

Table D.4: Subsample by year

	(1)	(2)	(3)	(4)	(5)	(6)
		1851-1881			1881-1911	
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.064***	0.038***	0.035***	0.115***	0.076**	0.070**
	(0.007)	(0.008)	(0.007)	(0.023)	(0.031)	(0.031)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	0.948***	0.790***	0.746***	2.128***	2.419***	2.328***
	(0.105)	(0.124)	(0.120)	(0.328)	(0.441)	(0.433)
Upward Mobility	0.034***	0.031***	0.029***	0.083***	0.097***	0.094***
	(0.005)	(0.006)	(0.006)	(0.013)	(0.019)	(0.019)
Downward Mobility	0.012***	0.010**	0.009**	0.034***	0.042***	0.040***
	(0.003)	(0.004)	(0.004)	(0.010)	(0.015)	(0.015)
Obs.		273,844			695,399	
County_{t-1} 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

Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, 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 HISCAM 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 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-1}. 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 and port (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 the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Figure D.6: Excluding one county at a time



Note: We estimate equation 1 excluding one county at a time. The figure plots the coefficient of the standardised Proximity$_{i,c,t-1}$, 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 childhood county$_{t-1}$. 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: Social Mobility Pattern by Urban-Rural

	(1)	(2)	(3)	(4)	(5)	(6)
		Urban			Rural	
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.101***	0.066***	0.064***	0.067***	0.051***	0.049***
	(0.030)	(0.022)	(0.022)	(0.009)	(0.011)	(0.011)
$\lvert HISCAM^{son} - HISCAM^{father}\rvert$	1.137***	0.916**	0.880**	1.145***	1.159***	1.130***
	(0.436)	(0.378)	(0.373)	(0.141)	(0.158)	(0.155)
Upward Mobility	0.066***	0.054***	0.053***	0.045***	0.052***	0.051***
	(0.021)	(0.017)	(0.017)	(0.006)	(0.007)	(0.007)
Downward Mobility	−0.006	0.003	0.002	0.012***	0.007	0.007
	(0.013)	(0.013)	(0.013)	(0.004)	(0.005)	(0.005)
Obs.		380,281			588,962	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County_{t-1} 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 $\text{Proximity}_{i,c,t-1}$, 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 HISCAM 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. Columns 1 to 3 (4 to 6) include the sample of sons who grew up in urban (rural) areas. All regressions include fixed effects for census year and childhood county_{t-1}. 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 and port (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 the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.6: Excluding individuals at nodes

	(1)	(2)	(3)
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.064***	0.039***	0.037***
	(0.011)	(0.013)	(0.012)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	1.038***	0.937***	0.911***
	(0.161)	(0.172)	(0.168)
Upward Mobility	0.048***	0.050***	0.050***
	(0.007)	(0.008)	(0.008)
Downward Mobility	0.004	−0.0004	−0.001
	(0.005)	(0.006)	(0.005)
SW-F	63.948	15.692	9.416
F-Stat	63.948	94.151	94.158
Obs.	813,513	813,513	813,513
County_{t-1} 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

Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, 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 HISCAM 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, 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-1}. 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 and port (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). SW-F is the Sanderson and Windmeijer (2016) F-statistic for weak instruments. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.7: Age of father

	(1)	(2)	(3)	(4)	(5)
Age of father	20-65	20-30	31-40	41-50	51-65
HISCO^{son} $\neq$ HISCO^{father}	0.071***	0.014	0.066***	0.069***	0.081***
	(0.011)	(0.039)	(0.013)	(0.012)	(0.012)
$\lvert HISCAM^{son} - HISCAM^{father}\rvert$	1.085***	0.540	1.021***	1.110***	1.101***
	(0.136)	(0.765)	(0.181)	(0.158)	(0.180)
Upward Mobility	0.047***	0.069*	0.062***	0.049***	0.034***
	(0.006)	(0.036)	(0.009)	(0.007)	(0.008)
Downward Mobility	0.009*	−0.037	−0.004	0.009	0.020***
	(0.005)	(0.037)	(0.007)	(0.006)	(0.006)
Obs.	969,243	6,068	221,143	451,323	290,709

Notes: Each entry represents the coefficient of the standardised Proximity$_{i,c,t-1}$, 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 HISCAM 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. Samples include father-son pairs where the father is aged 20-65 (column 1), 20-30 (column 2), 31-40 (column 3), 41-50 (column 4) and 51-65 (column 5). All regressions include fixed effects for census year and childhood county$_{t-1}$. 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. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.8: Age of son

	(1)	(2)	(3)	(4)	(5)
Age of son	10-22	10-11	12-13	14-16	17-22
HISCO^{son} $\neq$ HISCO^{father}	0.062***	0.052***	0.044***	0.073***	0.079***
	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	1.056***	0.879***	0.941***	1.279***	0.930***
	(0.144)	(0.184)	(0.207)	(0.210)	(0.190)
Upward Mobility	0.049***	0.048***	0.058***	0.054***	0.029***
	(0.006)	(0.008)	(0.010)	(0.009)	(0.009)
Downward Mobility	0.006	0.003	−0.006	0.010	0.018**
	(0.005)	(0.007)	(0.008)	(0.008)	(0.008)
Obs.	969,243	230,383	176,214	220,808	234,814

Notes: Each entry represents the coefficient of the standardised Proximity$_{i,c,t-1}$, 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 HISCAM 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. Sample includes father's age between the ages of 20-65 (column 1), 20-30 (column 2), 31-40 (column 3), 41-50 (column 4) and 51-65 (column 5). All regressions include fixed effects for census year and childhood county$_{t-1}$. 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. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.9: Social Mobility Pattern : Natives vs. Foreigners

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Natives			1st generation immigrant			2nd generation immigrant	
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.087***	0.063***	0.061***	−0.028	−0.074	−0.076	0.119***	0.089**	0.089**
	(0.010)	(0.010)	(0.010)	(0.045)	(0.084)	(0.085)	(0.028)	(0.036)	(0.036)
$\lvert HISCAM^{son} - HISCAM^{father}\rvert$	1.169***	1.022***	1.002***	0.776	1.017	1.016	1.735***	1.413*	1.413*
	(0.128)	(0.146)	(0.144)	(1.165)	(1.950)	(1.956)	(0.595)	(0.761)	(0.760)
Upward Mobility	0.049***	0.047***	0.046***	0.133*	0.234*	0.237**	0.069***	0.067**	0.066**
	(0.006)	(0.006)	(0.006)	(0.074)	(0.120)	(0.120)	(0.024)	(0.033)	(0.033)
Downward Mobility	0.011***	0.007	0.007	−0.113**	−0.161**	−0.164**	0.010	−0.001	−0.001
	(0.004)	(0.005)	(0.005)	(0.048)	(0.078)	(0.078)	(0.021)	(0.029)	(0.029)
Obs.		904,689			7,865			40,877	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County_{t-1} FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Historical travel routes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Each entry represents the coefficient of $\text{Proximity}_{i,c,t-1}$, 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 HISCAM 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. 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-1}. 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 and port (columns 2, 3, 5, 6, 8 and 9), and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (columns 3, 6 and 9). Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.10: Father as Stayer or Mover

	(1)	(2)	(3)	(4)	(5)	(6)
		Stayers			Movers	
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.036***	0.020**	0.016**	0.119***	0.091***	0.088***
	(0.007)	(0.008)	(0.008)	(0.015)	(0.015)	(0.015)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	0.743***	0.673***	0.617***	1.455***	1.264***	1.229***
	(0.117)	(0.143)	(0.141)	(0.180)	(0.193)	(0.190)
Upward Mobility	0.049***	0.052***	0.050***	0.046***	0.041***	0.039***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
Downward Mobility	−0.001	−0.005	−0.006	0.021***	0.016**	0.016**
	(0.004)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)
Obs.		405,743			563,500	
County_{t-1} 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

Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, 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 HISCAM 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. The sample include the sample of fathers who resided in their county of birth (columns 1 to 3) and father who haven't moved away (columns 4 to 6). All regressions include fixed effects for census year and childhood county_{t-1}. 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 and port (columns 2, 3, 5, 6, 8 and 9), and household characteristics consisting of the number of servants, household size and whether the father is born outside England and Wales (columns 3, 6 and 9). Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table D.11: Farming occupations

	(1)	(2)	(3)
	All	Farm	Non-farm
$\text{HISCO}^{son} \neq \text{HISCO}^{father}$	0.062***	0.118***	0.034**
	(0.010)	(0.016)	(0.014)
$\lvert HISCAM^{son} - HISCAM^{father} \rvert$	1.056***	1.237***	0.659***
	(0.144)	(0.198)	(0.168)
Upward Mobility	0.049***	0.024***	0.022***
	(0.006)	(0.007)	(0.007)
Downward Mobility	0.006	0.029***	0.010
	(0.005)	(0.008)	(0.006)
Obs.	969,243	226,466	742,777

Notes: Each entry represents the coefficient of the standardised $\text{Proximity}_{i,c,t-1}$, 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 HISCAM 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 (column 1). Sample is restricted to fathers in farming (column 2) and non-farming (column 3). All regressions include fixed effects for census year and childhood county_{t-1}. 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. Standard errors clustered at the parish in year $t-1$ are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.